

中国人口分布规律及演化机理研究^{*}

施华萍 柯见洪[†] 孙 策 林振权

(温州大学物理与电子信息学院 温州 325035)

(2008 年 5 月 14 日收到, 2008 年 7 月 10 日收到修改稿)

近 20 年来, 中国的总人口近似地按指数规律增加, 但增加的速度有减缓的趋势. 省、市、县三级的人口分布具有相似的规律: 人口数 k 小的各级人口分布随 k 缓慢地变化, 而 k 较大时, 人口分布随 k 急剧变化, 满足齐普夫定律. 研究结果表明中国人口系统的各级分布具有自相似性. 采用无标度网络上的聚集体迁移模型较好地解释了中国人口分布的演化规律.

关键词: 人口分布, 齐普夫定律, 迁移, 动力学行为

PACC: 0175, 0565

1. 引 言

人口迁移与城市化是社会经济总体发展的一部分, 也是当前包括我国在内的许多发展中国家面临的一个重要的问题. 我国是一个人口大国, 人口问题是关系我国经济社会能否持续健康发展的关键性问题. 所以, 我国人口的演化动力学一直是社会学家关注的一个热点课题. 前人在研究人口问题时, 一般运用人口学或社会学的方法宏观上分析城镇人口分布的规律. 近年来, 以 Stanley 和 Redner 为代表的研究人员基于非平衡统计物理思想研究宏观经济社会现象的微观动力学演化机理, 取得重大进展^[1-11]. 物理学家采用动力学模型的方法来研究经济社会系统的复杂性, 不仅拓宽了经济社会学课题的研究思路, 也能够从微观上解释一些宏观的经济社会系统的内在的演化机理.

经济社会的一些系统, 例如具有一定人口的城镇、某一行业具有一定规模的企业等, 都可以看成是广义的聚集体. 聚集体非平衡演化是自然界中的一种普遍现象, 例如, 星云聚集成恒星, 动物种群的演化, 血管中红血球的聚集, 带电离子的聚集等等^[12]. 近 20 多年来, 物理学家开展了大量的研究工作, 深入地探讨了聚集体非平衡演化动力学^[13-17]. 在研究自然科学领域的聚集过程动力学中发展起来的一些

研究方法可以被用于研究经济社会系统的演化过程^[18-21]. Stanley 等研究了美国同行业内的企业规模, 发现其增长率分布满足标度律^[1-3]. Redner 和 Krapivsky 等提出了单体迁移(交换)驱使聚集体生长模型, 分别研究具有相互作用的个体财富分布和城镇人口分布等的演化动力学标度行为^[4, 5]. 最近, 我们采用非平衡统计物理方法对经济社会系统的演化行为进行了探索, 研究了交换驱使的聚集生长, 以及交换与聚合、自产生与消亡、催化产生等机理共同驱使的聚集演化过程的动力学行为^[18, 19]. 我们还在无标度网络上研究了迁移驱使的聚集体生长行为, 并利用该动力学模型定量地解释了 1900—2000 年间美国各郡的人口分布^[20, 21].

本文在分析我国人口实际分布规律的基础上, 建立合理的人口演化模型(例如将人口的出生率与死亡率考虑在内), 然后从理论上分析人口分布的规律.

2. 我国人口分布规律

2.1. 总人口增长规律

我们采用 $M(t)$ 来表示 t 时刻我国总人口. 图 1 描述了 1980—2004 年间我国总人口随时间的增加规律, 其中横坐标 t 表示年份, 纵坐标

^{*} 国家自然科学基金(批准号: 10775104, 10305009)资助的课题

[†] 通讯联系人. E-mail: kejianhong@yahoo.com.cn

$\ln[M_1(t)/M_1(1980)]$ 表示 t 时刻我国总人口与 1980 年总人口的比值的对数值. 图 1 中的点表示真实的人口数据, 而线表示人口数据的线性拟合曲线, 其斜率为 0.0140 ± 0.0003 . 从图 1 可以看出, 我国的总人口近似地按指数规律增加, 即 $M_1(t) \propto e^{\alpha t}$. 图 2 按时间段分析了 $\ln[M_1(t)/M_1(1980)]$ 随时间的变化规律, 进一步说明我国总人口随时间指数增加. 但是我国的人口增加速率从上世纪 90 年代中期以来已经有所减小, 由图 2 可见, 1980—1994 年间 α 值近似为 0.0157 ± 0.0002 , 而 1995—2004 年间 α 值减小为 0.0102 ± 0.0003 , 而且 α 值将可能随时间进一步减小.

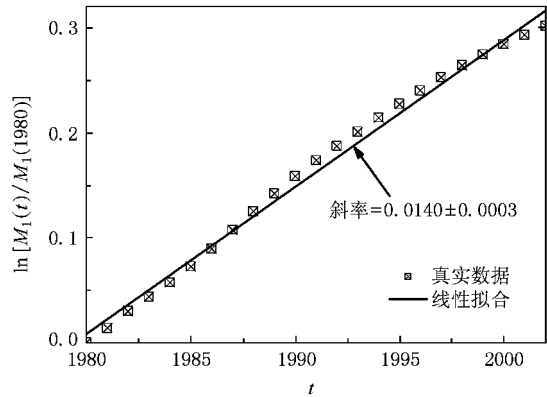


图 1 1980—2004 年间我国总人口 $M_1(t)$ 随时间 t 的演化

2.2. 省人口分布规律

我们分析了我国近 20 年的人口数据^[22—26]. 由

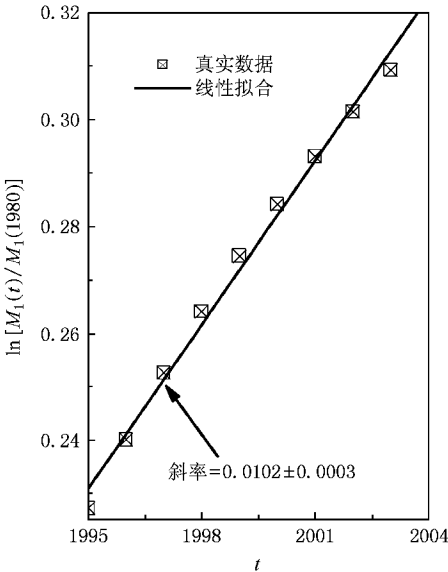
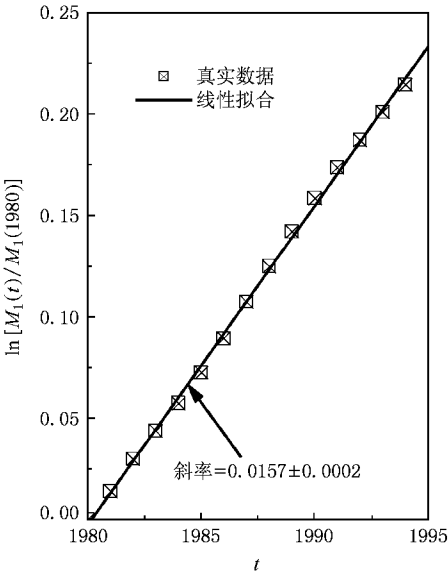


图 2 我国总人口 $M_1(t)$ 随时间 t 的演化在不同时间段的比较

于各省(或县、市)的实际人口数是一组离散的数据, 前人一般采用柱形图来描述人口分布. 柱形图能够直观地提供一些人口分布的信息, 例如在某一人口数区域的省的个数, 但利用柱形图只能粗略地分析人口分布随人口数的变化规律. 本文采用累积分布来统计分析我国人口分布规律. 引入累积统计量 $P_S(k, t)$ 来表示 t 时刻人口数大于或等于 k 的省(包括直辖市)的个数与我国省的总个数的比值(即人口累积分布). $P_S(k, t)$ 对人口数 k 的一阶导数就是我国的省人口分布函数. 此外, 因为不同年份下全国的总人口数是不一样的, 为了更合理地比较不同年份的省人口分布规律, 我们有必要消除总人口数的影响. 这里, 我们采用与文献^[20]一致的数据统计方法, 引入了总人口因子 $R(t) = M_1(t)/M_1(1986)$, 并将 $P_S(k, t)/R(t)$ 做为 t 时刻修正的省人口累积分布. 图 3 是不同年份省为统计单位修正的人口累积分布曲线. 由图 3 可以看出, 不同年份的省人口累积分布几乎具有相同的规律. 人口数 k 较小时, 省人口累积分布 $P_S(k, t)/R(t)$ 随着 k 变化比较缓慢, 当 k 达到一定大小后, $P_S(k, t)/R(t)$ 下降得很快. 此外, 由图 3 可以发现 k 大的省人口累积分布近似满足幂律分布 $P_S(k, t)/R(t) \sim k^{-\gamma}$, 而且不同年份的指数 γ 都近似为 1.34 ± 0.08 . 这说明, 我国以省为统计单位的人口分布服从齐普夫(Zipf)定律^[27], 其幂律分布指数约为 2.34. 这个现象在其他国家也普遍存在^[28, 29].

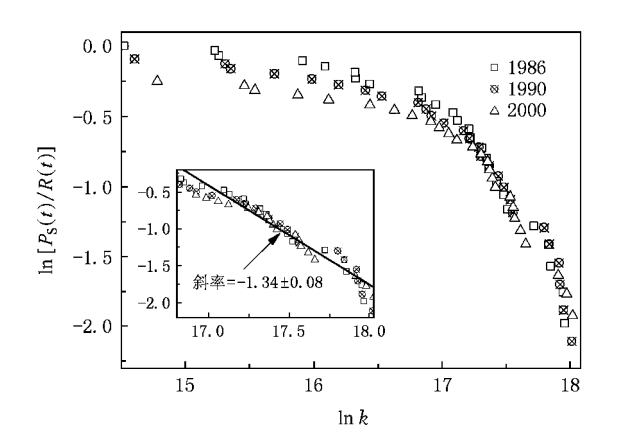


图3 以省为统计单位修正的人口累积分布 $P_S(k, t) / R(t)$ 的数据图

2.3. 市人口分布规律

图4所示的是以地级市为统计单位的修正的人口累积分布,其中 $C_S(k, t)$ 表示 t 时刻人口数大于或等于 k 的地级市的个数与我国地级市的总个数的比值.由图4可见,人口数 k 较小时,市人口累积分布随着 k 的增加缓慢减小;当 k 很大时,市人口累积分布随 k 急剧下降,并近似服从幂律分布.表1列出了不同年份的市累积人口分布的幂律分布指数 γ ,其中 $4800000 \leq k \leq 9000000$.

我们还进一步分析了人口数 k 较小时市人口

累积分布的规律.由图5可见,当 $k < 1500000$ 时, $C_S(k, t) / R(t)$ 与 $k^{0.9 \pm 0.1}$ 近似地呈线性关系.

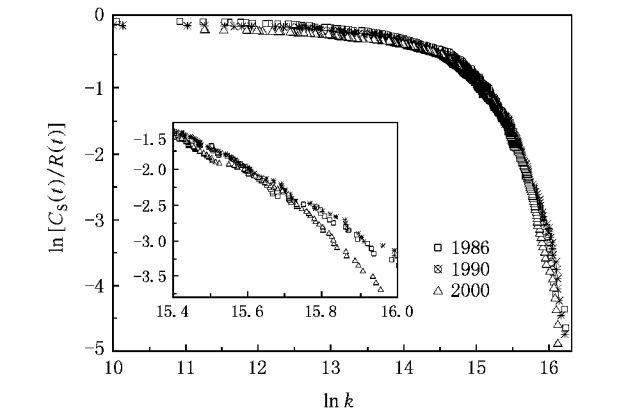


图4 以地级市为统计单位的修正的人口累积分布 $C_S(k, t) / R(t)$ 的数据图

表1 市修正人口累积分布 $C_S(k, t) / R(t)$ 对 k 的幂律分布指数 γ		
时间	斜率	误差
1986 年	3.41	0.07
1990 年	3.39	0.06
2000 年	3.97	0.09

2.4. 县人口分布规律

图6是以县(包含区)为统计单位的修正的人口

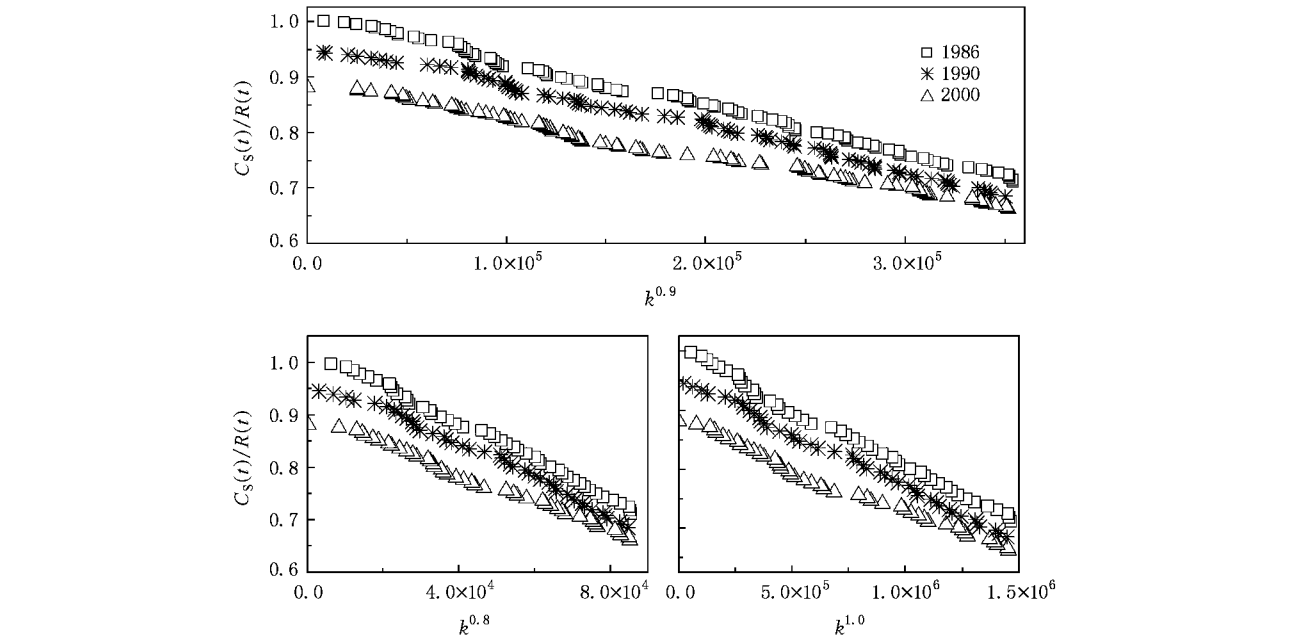


图5 人口数 k 较小的地级市人口累积分布

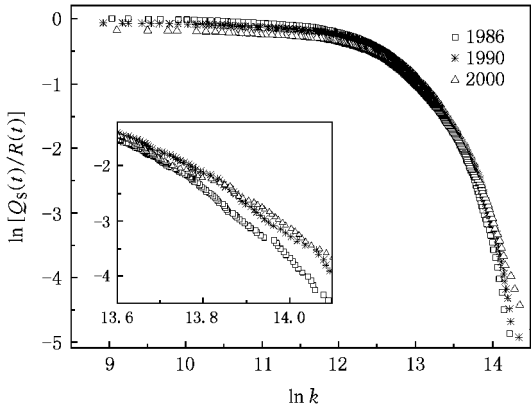


图 6 以县为统计单位的修正的人口累积分布 $Q_s(k,t)/R(t)$ 的数据图

累积分布,其中 $Q_s(k,t)$ 表示 t 时刻人口数大于或等于 k 的县的个数与我国县的总个数的比值.与省、市两级的人口累积分布相似,人口数 k 较小时,县

人口累积分布随着 k 的增加缓慢减小;当 k 很大时,县人口累积分布随 k 急剧下降,并近似服从幂律分布.表 2 给出了不同年份的县人口累积分布的幂律分布指数 γ ,其中 $800000 \leq k \leq 1300000$. 这些结果表明人口数较大的省、市或县的人口分布都服从齐普夫定律,不过其幂律分布指数有所不同.

表 2 县修正人口累积分布 $Q_s(k,t)/R(t)$ 对 k 的幂律分布指数 γ

时间	斜率	误差
1986 年	3.75	0.02
1990 年	3.19	0.03
2000 年	2.80	0.02

我们也进一步分析了人口数 k 较小时县人口累积分布的规律(图 7).由图 7 可见,当 $k < 230000$ 时, $Q_s(k,t)/R(t)$ 与 $k^{1.1 \pm 0.1}$ 近似地呈线性关系.

由图 3—7 可见,我国的人口分布系统具有一定的自相似性.

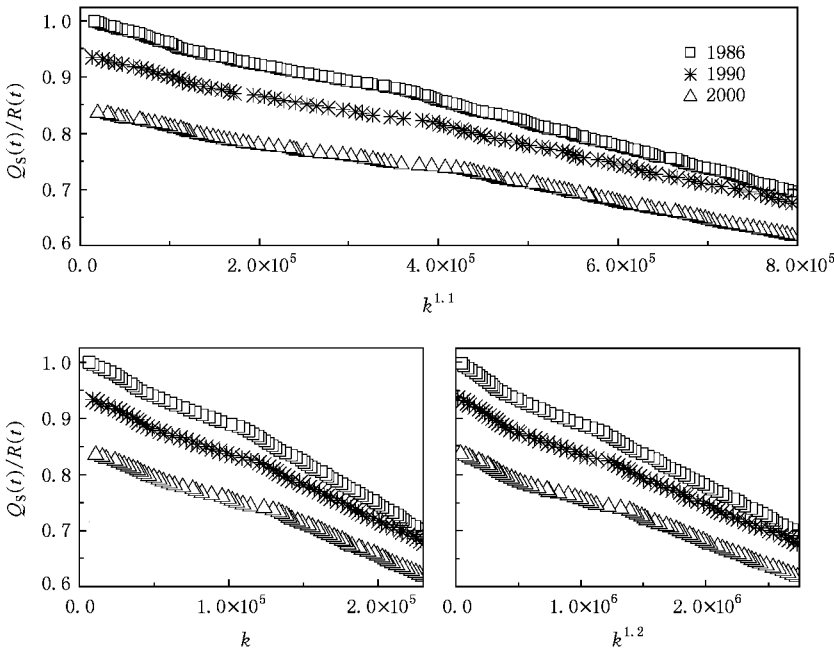


图 7 人口数 k 较小的县人口累积分布

3. 人口系统的演化机理

3.1. 人口系统的演化模型

上述的数据图说明,我国县(市)人口分布与美国城市人口分布有着相似的演化规律^[20].因此,我们采用复杂网络上的迁移模型来解释我国的人口分

布.基于复杂网络的人口迁移机理可以由如下反应式来描述:

$$A_{k,l} + A_{i,j} \xrightarrow{K(k;l|i;j)} A_{k-1,l} + A_{i+1,j},$$

其中 $A_{k,l}$ 表示占据在一个具有 l 个连接(也称度)的节点上、大小为 k 的聚集体, $K(k;l|i;j)$ 表示一个单体从聚集体 $A_{k,l}$ 迁移到 $A_{i,j}$ 的速率核.对应一个人口系统,节点代表县(市),聚集体代表人口集团;节点间的连线代表两个县(市)的交通连接,聚集体的

大小代表该县(市)的人口数.事实上,一个国家内的人口迁移总是通过交通网络(复杂网络)来实现.此外,一个国家的人口分布还与人口的出生与死亡有密切的关系.在文献[20]的基础上,本文进一步考虑了聚集体自出生与自死亡.聚集体自出生可表示为反应式 $A_{k,l} \xrightarrow{K(k;l)} A_{k+1,l}$,聚集体自死亡可表示为反应式 $A_{k,l} \xrightarrow{K(k;l)} A_{k-1,l}$,其中 $K(k;l)$ 和 $K(k;l)$ 分别是个体出生与死亡的速率核.

我们用速率方程的方法研究了复杂网络上聚集体的迁移动力学. t 时刻,占据在连接度为 l 的节点上、大小为 k 的聚集体 $A_{k,l}$ 的浓度设为 $a_{kl}(t)$,即聚集体 $A_{k,l}$ 的个数与所有聚集体总个数的比值.假设在反应过程中,反应速率与参与反应的每个反应物的浓度成正比.我们得到了本文所研究的模型的速率方程

$$\begin{aligned} \frac{da_{kl}}{dt} = & a_{k+1,l} \sum_{i,j} K(k+1;l|i;j) a_{ij} \\ & + a_{k-1,l} \sum_{i,j} K(i;j|k-1;l) a_{ij} \\ & - a_{kl} \sum_{i,j} [K(k;l|i;j) + K(i;j|k;l)] a_{ij} \\ & + K(k-1;l) a_{k-1,l} + K(k+1;l) a_{k+1,l} \\ & - [K(k;l) + K(k;l)] a_{kl}. \end{aligned} \quad (1)$$

(1)式的边界条件为 $a_j(t) \equiv 0$,其中 $j = -1, -2, -3, \dots$ 和 $l = 1, 2, 3, \dots$. (1)式中,右边第一项描述一个单体从聚集体 $A_{k+1,l}$ 迁出到其他聚集体,则该聚集体将变成聚集体 $A_{k,l}$,从而增加了 $A_{k,l}$ 的浓度;第二项描述一个单体由其他聚集体迁入到聚集体 $A_{k-1,l}$,则该聚集体将变成聚集体 $A_{k,l}$,从而增加了 $A_{k,l}$ 的浓度.相似地,第三项描述一个单体从聚集体 $A_{k,l}$ 迁出到其他聚集体,而第四项描述一个单体由其他聚集体迁入到聚集体 $A_{k,l}$,从而都导致了 $A_{k,l}$ 的浓度的减小.(1)式右边的第五项描述聚集体 $A_{k-1,l}$ 自发产生一个新单体而成为 $A_{k,l}$,从而增加 $A_{k,l}$ 的浓度;第六项描述聚集体 $A_{k+1,l}$ 由于一个单体死亡而变成 $A_{k,l}$,从而增加 $A_{k,l}$ 的浓度;第七和第八项分别描述聚集体 $A_{k,l}$ 自发产生或死亡一个单体,从而导致了 $A_{k,l}$ 浓度的减小.

3.2. 模型的解析解

(1)式实际上是一组无数多个耦合的非线性微分方程,在一般情况下很难得到它的解析解.为了得

到(1)式的显解并由此来拟合我国的实际人口分布曲线,我们选取了一个简单的迁移速率核 $K(k;l|i;j) = K(ki)^u(lj)^v$,其中 u 是与聚集体大小有关的指数, v 是与复杂网络节点的度值有关的指数(为简化起见,可以设 $K = 1$).文献[20]也采用了相同的迁移速率核来研究无标度网络上的迁移动力学.当 u 增大时,聚集体中的单体容易迁出,同时也容易吸引新的单体迁入;当 v 增大时,驻扎在度值大的节点上的聚集体更容易发生单体的迁出或迁入.对应一个人口系统,指数 u 描述一个县(市)居民的“不安分”程度(即迁移意愿),而指数 v 描述人口迁移对交通路线的依赖程度.一般而言,人口迁移系统的 u 和 v 都是正数,因为人口众多和交通方便的大城市与那些交通不便的小城市相比,人口更容易迁出,但也能吸引更多的人口迁入.此外,我们假定一个聚集体中的每个单体有相同的概率生产出一个新单体或自死亡,那么一个聚集体的单体出生率和死亡率就与该聚集体含有的单体数成正比,即 $K(k;l) = lk$ 和 $K(k;l) = Jk$ (l 和 J 是常数).对于人口系统,选择这样的出生/死亡速率核应该是合理的.

为了便于求解,我们考虑一个简单的初始条件:在 $t = 0$ 时系统中只存在单体.另一方面,已有的研究表明,交通网络具有无标度的拓扑结构^[30-31].这里,我们假定聚集体所占据的节点的度值分布具有无标度的幂律形式 $P(l) = l^{-\gamma}$.那么系统的初始条件可以表示成

$$a_{kl}(0) = (\gamma - 1)l^{-\gamma}\delta_{kl}. \quad (2)$$

在(2)式中,引入 $\gamma - 1$ 是为了归一化.此外,我们也必须指出,对于一个真实的人口系统,节点的连接度值 l 应该存在一个最大的截止值 l_{\max} .但是,通过分析速率方程(1)可知,连接度值 l 的取值范围对 $a_{kl}(t)$ 的形式几乎没有影响^[20-21].为了计算方便,我们假设连接度值的取值可以从 1 到 ∞ 的任意整数.

基于上面的一些假定,(1)式可以简写为

$$\begin{aligned} \frac{da_{kl}}{dt} = & l^v [(k+1)^u a_{k+1,l} + (k-1)^u a_{k-1,l} \\ & - 2k^u a_{kl}] \sum_{i,j} i^u j^v a_{ij} \\ & + K(k-1) a_{k-1,l} - lka_{kl} \\ & + K(k+1) a_{k+1,l} - Jka_{kl}. \end{aligned} \quad (3)$$

将(3)式两边同时乘以 k ,然后对所有 k 与 l 相加,便可得到 $\dot{M}_l(t) = \sum_{k,l} k \dot{a}_{kl}(t) = (I - J)M_l(t)$,它的解为 $M_l(t) = e^{(I-J)t}$.令 $\alpha = I - J$.很明显, α 代

表人口的净出生率. 图 1 表明我国的净出生率大于 0.

在单体分布的初始条件下, 可以假定 (3) 式的解 $a_{kl}(t)$ 满足标度形式

$$a_{kl}(t) = e^{at} l^{-\lambda} k^{-\tau} S^{\tau-2} \Phi(k/S), \quad (4)$$

其中 $S(t, l)$ 表示位于连接度值为 l 的节点上的聚集体的特征大小. (4) 式中的 $S(t, l)$ 与临界现象的关联长度有相类似的物理含义. 这样问题变为求标度函数 $\Phi(x)$, 特征尺寸 $S(t, l)$, 以及标度指数 λ 和 τ .

令 $M_{lu}(t) = \sum_{k=1}^{\infty} k^u a_{kl}(t)$. 利用 (3) 式, 可以得到

$$\frac{dM_{l0}}{dt} = -a_{l1} \left(l^v \sum_{k,l} k^u l^v a_{kl} + J \right), \quad (5)$$

$$\begin{aligned} \frac{dM_{l2}}{dt} &= 2M_{lu} l^v \sum_{k,l} k^u l^v a_{kl} \\ &+ 2\alpha M_{l2} + (I + J)M_{l1}. \end{aligned} \quad (6)$$

将 (4) 式分别代入 (5) 和 (6) 式, 可以推导出 $\tau = u - 1$ 以及

$$S(t, l) \sim e^{at(3-2u)} l^{v(2-u)}. \quad (7)$$

由 (7) 式可知, 标度假设 (4) 成立的前提是 $u < 3/2$. 假设 $u > 3/2$ 的情况下标度假设 (4) 式也是正确的, 我们将推导出自相矛盾的解 $S(t, l) \propto [(u-2)\chi t_c - e^{at}]^{\chi(2-u)}$, 这里 t_c 是一个有限的正值. 在 $u > 3/2$ 的情况下, 系统将会发生类似凝胶的相变. 对于特殊情况 $u = 3/2$, 聚集体大小分布只符合修正的标度形式^[21]. 在本文中, 只讨论 $u < 3/2$ 的情况.

利用 (3) 式, 可以得到 $\sum_{k=0}^{\infty} \dot{a}_{kl}(t) = 0$, 它的解为 $\sum_{k=0}^{\infty} a_{kl}(t) \equiv (\gamma - 1)l^{-\gamma}$. 因为 $t \rightarrow \infty$ 时 $M_{l0}(t) \rightarrow 0$, 所以 $a_{0l}(t) \sim l^{-\gamma}$. 将 (4) 式代入 (3) 式, 便可以得到 $\lambda = \gamma - v(2 - u)$.

下面我们求标度函数 $\Phi(x)$. 将 (4) 式代入 (3) 式, 可以得到标度函数 $\Phi(x)$ 的微分方程

$$\begin{aligned} x\Phi''(x) + (2 + px^{2-u})\Phi'(x) \\ = p(u-3)x^{1-u}\Phi(x), \end{aligned} \quad (8)$$

其中 p 是一个分离常数, 它满足关系 $dS(t, l)/dt = p[S(t, l)]^{u-1}$. 求解 (8) 式可得

$$\Phi(x) = C_2 \exp(-C_1 x^{2-u}), \quad (9)$$

其中 $C_1 = p(2-u)$, C_2 是一个积分常数. C_2 与 p 是相关联的, 可以选择一个合适的 p 值, 使 $C_2 = 1$.

由此我们可以得到 $u < 3/2$ 时聚集体大小分布 $a_{kl}(t)$ 的解析解

$$\begin{aligned} a_{kl}(t) &= C_4 l^{-\gamma-v} k^{1-u} e^{-uat(3-2u)} \exp(-C_1 x^{2-u}), \\ x &= C_3 k e^{-at(3-2u)} l^{-v(2-u)}, \end{aligned} \quad (10)$$

其中 $C_4 = C_3^{u-3}$ (C_3 是一个积分常数). 显然, 对于给定 l , (10) 式可以写成标度形式

$$a_{kl}(T) \sim k^{-\tau} T^{-w+1} \Phi[k/S(T)] \quad S(T) \propto T, \quad (11)$$

其中 $\tau(t) = e^{at}$. 标度指数有如下的关系: $\tau = u - 1$, $z = 1/(3-2u)$ 和 $w = (2-\tau)z$. 将上述的结果与文献 [21] 中的结果进行比较, 可以发现伴随着聚集体自发生/死亡的聚集体迁移系统与单纯的聚集体迁移系统有着相似的演化行为, 它们的聚集体大小分布都满足相似的标度形式, 而且标度指数和标度函数完全相同. 这些结果表明, 在选择上述速率核的情况下, 聚集体自发生/死亡对聚集体大小分布的标度形式并没有影响, 聚集体大小分布的演化本质上由聚集体间的迁移所决定.

下面求具有同样大小 k 的所有聚集体的累积分布, 即 $c_k(t) = \sum_{l=1}^{\infty} a_{kl}(t)$. 当 $u < 3/2$ 时, 可以得到

$$c_k(t) \approx k^{-\xi} e^{-\zeta t} \Psi(y), \quad y = C_3 k e^{-at(3-2u)}, \quad (12)$$

其中 $\xi = 1 + (\gamma - 1)(2 - u)/v$, $\zeta = [1 - (\gamma - 1)(2 - u)/v](3 - 2u) - 1$, 而标度函数为

$$\Psi(y) = C_5 \int_0^{C_1 y^{2-u}} Z^{\gamma-1} y^v \exp(-Z) dZ, \quad (13)$$

其中 $C_5 = C_4 v^{-1} (C_1 C_3^{2-u})^{-(\gamma+v-1)/v}$. 对于 $1 \ll k \ll e^{at(3-2u)}$, 由 (13) 式可推导出

$$c_k(t) \sim k^{1-u} e^{-uat(3-2u)}. \quad (14)$$

而对于 $k \gg e^{at(3-2u)}$, $c_k(t)$ 满足

$$c_k(t) \sim k^{-\xi} e^{-\zeta t}. \quad (15)$$

(14) 和 (15) 式表明, 当 k 较小时, 聚集体大小累积分布 $c_k \sim k^{1-u}$; 当 k 很大时, $c_k \sim k^{-\xi}$.

3.3. 解析结果拟合实际人口分布曲线

(14) 和 (15) 式与我国省、市、县三级人口分布定性上都相似 (见图 3.4.6). 由此可见我们的模型能够从微观机理上定性地解释我国的人口分布规律. 下面以我国的县人口分布为例, 利用聚集体大小分布的解析解 (10) 来定量地拟合真实的人口分布曲线.

由图 7 可知, 对于小 k , $Q_S(k, t)$ 与 $k^{1.1 \pm 0.1}$ 呈线性关系. 而 $Q_S(k, t) = \sum_{j \geq k} c_j(t)$, 于是可得 $c_k(t) \sim k^{0.1 \pm 0.1}$. 因此, 在 (10) 式中, $u = 0.9 \pm 0.1$ 指数 γ

和 v 的关系是 $\xi = 1.0 + 1.1(\gamma - 1)v$. 在经过多次尝试后发现,拟合结果强烈地取决于 ξ 的值,而在 $2 < \gamma < 3$ 之间即使选择不同的 γ 值,可以得出相似的结果.取 $\gamma = 2.5$,我们得到

$$Q_s(k,t) = C_6 \int_k^\infty dx x^{0.1} \int_1^\infty dy y^{-2.5-1.65(\xi-1.0)} \times \exp[-C_7 x^{1.1} y^{-1.65(\xi-1.0)}], \quad (16)$$

其中 C_6 和 C_7 是常数. C_6, C_7 和 ξ 的值可以通过分析实际的人口数据求得.

利用(16)式我们绘制出县人口累积分布的拟合曲线,并与真实数据相比较,如图8所示.由图8可见,我们的理论计算结果较好地拟合实际人口分布数据.

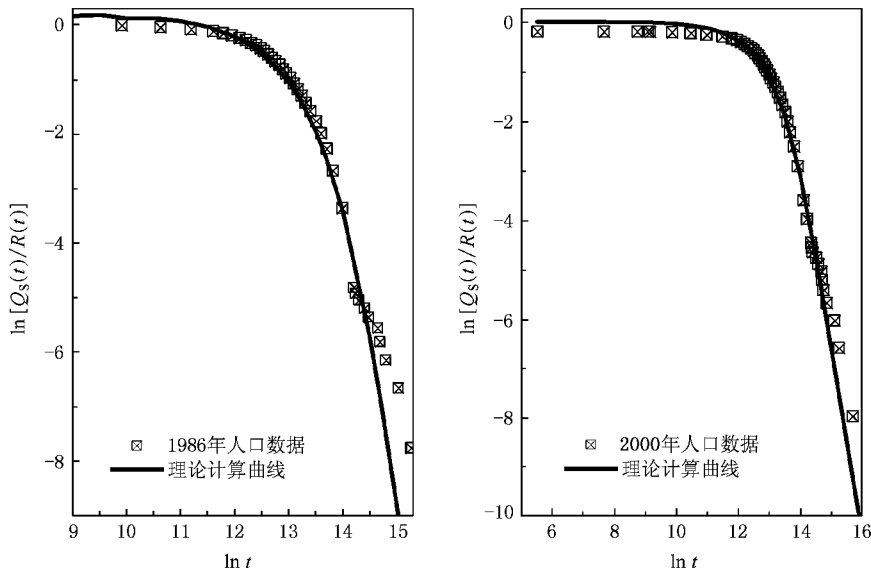


图8 县人口累积分布 $Q_s(k,t)/R(t)$ 的理论计算结果与真实数据的比较

4. 结 论

本文研究了最近 20 年来我国人口分布的特点.从总体上来讲,我国的人口总量一直处于增长过程,但是自上世纪 90 年代后人口增长的趋势已经有所减缓.同时,分别给出了省、市、县三级的不同年份的

人口分布曲线,发现我国的人口系统的分布具有一定的自相似性.当 k 较小,县(市)人口分布在各年的曲线斜率大小很接近.当 k 较大时,县(市)人口分布满足齐普夫定律.同时,我们采用无标度网络上的聚集体迁移模型来模拟我国的人口演化,理论计算结果与真实的人口数据符合得比较好.

[1] Stanley M H R , Amaral L A N , Buldyrev S V , Havlin S , Leschhorn H , Maass P , Salinger M A , Stanley H E 1996 *Nature* **379** 804

[2] Plerou V , Amaral L A N , Gopikrishnan P , Meyer M , Stanley H E 1999 *Nature* **400** 433

[3] Fu D , Panmooli F , Buldyrev S V , Riccaboni M , Matia K , Yamasaki K , Stanley H E 2005 *Proc. Natl. Acad. Sci.* **102** 18801

[4] Ispolatov S , Krapivsky P L , Redner S 1998 *Eur. Phys. J. B* **2** 267

[5] Leyvraz F , Redner S 2002 *Phys. Rev. Lett.* **88** 068301

[6] Ben-Naim E , Vazquez F , Redner S 2006 *Eur. Phys. J. B* **49** 531

[7] Amaral L A N , Gopikrishnan P , Plerou V , Stanley H E 2001 *Physica A* **299** 127

[8] Zhang P P , He Y , Zhou T , Su B B , Chang H , Zhou Y P , Wang B H , He D R 2006 *Acta Phys. Sin.* **55** 60 (in Chinese) [张培培、何 阅、周 涛、苏蓓蓓、常 慧、周月平、汪秉宏、何大韧 2006 物理学报 **55** 60]

[9] Guan J Y , Wu Z X , Wang Y H 2007 *Chin. Phys.* **16** 3566

[10] Lin H , Wu C X 2007 *Acta Phys. Sin.* **56** 4313 (in Chinese) [林海、吴晨旭 2007 物理学报 **56** 4313]

[11] Xu B Z 2008 *Acta Phys. Sin.* **57** 1991 (in Chinese) [徐炳振 2008 物理学报 **57** 1991]

- [12] Friedlander S K 1977 *Smoke , dust and haze : fundamental of aerosol behavior* (New York : Wiley)
- [13] Meakin P 1983 *Phys. Rev. Lett.* **51** 1119
- [14] Vicsek T , Family F 1984 *Phys. Rev. Lett.* **52** 1669
- [15] Ziff R M , Hendriks E M , Ernst M H 1985 *Phys. Rev. Lett.* **49** 593
- [16] Yu J , Hu G 1989 *Phys. Rev. B* **40** 661
- [17] Ben-Naim E , Krapivsky P L 2003 *Phys. Rev. E* **68** 031104
- [18] Lin Z , Ke J , Ye G 2006 *Phys. Rev. E* **74** 046113
- [19] Wang H , Lin Z , Ke J 2007 *Phys. Rev. E* **75** 046108
- [20] Ke J , Lin Z , Zheng Y , Chen X , Lu W 2006 *Phys. Rev. Lett.* **97** 028301
- [21] Ke J , Chen X , Lin Z , Zheng Y , Lu W 2006 *Phys. Rev. E* **74** 056102
- [22] China Population Information Centre (Eds.) 1986 *The handbook of China population data of 1986* (Beijing : China Population Information Centre)(in Chinese)[中国人口情报中心编 1986 中国人口资料手册 (北京 : 中国人口情报中心)]
- [23] China Population Information Centre (Eds.) 1987 *The handbook of China population data of 1987* (Beijing : China Population Information Centre)(in Chinese)[中国人口情报中心编 1987 中国人口资料手册 (北京 : 中国人口情报中心)]
- [24] National Bureau of Statistics of China (Eds.) 1991 *China Population Statistics Yearbook of 1990* (Beijing : China Statistics Press)(in Chinese)[国家统计局统计司编 1991 中国人口统计年鉴 1990 (北京 : 中国统计出版社)]
- [25] National Bureau of Statistics of China (Eds.) 2000 *China Population Statistics Yearbook of 2000* (Beijing : China Statistics Press)(in Chinese)[国家统计局统计司编 2000 中国人口统计年鉴 2000 (北京 : 中国统计出版社)]
- [26] National Bureau of Statistics of China (Eds.) 2004 *China Population Statistics Yearbook of 2004* (Beijing : China Statistics Press)(in Chinese)[国家统计局统计司编 2004 中国人口统计年鉴 2004 (北京 : 中国统计出版社)]
- [27] Zipf G K 1949 *Human Behavior and the Principle of Least Effort* (Cambridge : Addison-Wesley Press)
- [28] Moura N J Jr. , Ribeiro M B 2006 *Physica A* **367** 441
- [29] Newman M E J 2005 *Contemp. Phys.* **46** 323
- [30] Li W , Cai X 2004 *Phys. Rev. E* **69** 046106
- [31] Guimera R , Amaral L A N 2004 *Eur. Phys. J. B* **38** 381

Rules of the population distribution of China and its evolution mechanism ^{*}

Shi Hua-Ping Ke Jian-Hong[†] Sun Ce Lin Zhen-Quan

(College of Physics and Electronic Information , Wenzhou University , Wenzhou 325035 , China)

(Received 14 May 2008 ; revised manuscript received 10 July 2008)

Abstract

The characteristics of the population distribution of China is analyzed based on the method of statistical physics. The results show that in the last 20 years , the total population of China increases roughly exponentially with time and the growth rate is found to be slowing down. The population distribution in provinces , cities or counties follows approximately the same rule : the population distribution changes slowly for small k , while it changes rapidly for large k and the distribution obeys the Zipf 's law. This shows that the population distribution of China has self-similarity. A solvable model of migration-driven aggregate growth based on scale-free network is also proposed to simulate the population size distribution of China. It is found that the theoretical results are in good agreement with the realistic data.

Keywords : population distribution , Zipf 's law , migration , kinetic behaviour

PACC : 0175 , 0565

^{*} Project supported by the National Natural Science Foundation of China (Grant Nos. 10775104 and 10305009).

[†] Corresponding author. E-mail : kejianhong@yahoo.com.cn