

中国省级人口分布影响因素 的定量分析^{*}

孟向京 贾绍凤

(中国科学院 地理研究所)

提 要: 本文主要用1987年的数据,对影响省级人口密度的31种因素指标进行了相关和回归分析,建立了各种因素对人口密度的回归公式。作者引进了两个新的经济指标:单位国土面积粮食产量和国民收入。分析表明这两个指标比其他经济指标与人口密度的相关更紧密,自然因素尤其是海拔影响最大,其次是经济因素。

主题词: 人口分布 影响因素 相关分析

人口分布的影响因素是复杂多样的,这些因素包括了自然、经济、社会、政治、科教、文化、历史等各个方面。过去已有很多有关人口分布影响因素的研究成果,其中尤以人口地理学者贡献最大^[1]。

但总的来看,以往的研究成果主要是侧重于对人口分布的影响因素进行定性分析,而这些因素对我国人口分布的影响程度分别有多大,究竟哪些因素对我国的人口分布起主导作用等问题则无法作出满意的解答。本文尝试对人口分布的影响因素进行进一步的定量分析,以期对以上问题有更深入的了解和认识。并为人口分布的预测、调控等提供一定的科学依据。

1 指标的选取及分析

本文主要用相关分析的方法对影响人口分布的各因素进行定量分析。在选取人口分布的定量分析因素指标时,首先要基于以往定性分析的成果,以人口分布理论为指导,其次要考虑因素的量化及数据的可获取性。据此,我们选取了自然、经济、人口、教育四类共31种影响因素指标(表1)。其中自然因素指标12种,经济因素指标11种,人口因素指标6种,教育因素指标2种。对于某种人口分布的影响因素,可以用多个数量指标来表示,例如海拔是影响人口分布的重要因素,它可以用平均海拔、海拔100m、300m、和海拔500m以下的国土面积占总面积的百分比等指标表示。为了分析比较,对一种因素的多种指标也进行了分析。

1.1 单因素相关回归分析

本文1992年5月12日收到,1993年11月25日收到修改稿。

^{*} 本文得到国家自然科学基金资助。

我们主要用1987年的人口密度和以上31种指标的分省数据¹⁾，分析各种因素与人口密度的相关关系。

各单项因素指标与人口密度的相关分析结果列在表2中，表中同时给出了单因素回归公式，并按相关系数的大小进行排序。由于上海具有特殊性，其数据与其他省市偏差较大，在进行一些因素的相关分析时舍弃了上海市的数据。另外海南省、台湾省缺乏数据资料，所以分析时未能包括在内。

在0.05检测水平下，数据系列长度为29（即有29个省市自治区的数据），显著相关的临界指标为相关系数大于0.355；在0.01水平下，相应的相关系数应大于0.456。根据这一标准，从表2可知，在0.05检测水平下，在所分析的31种指标中，有22种指标与人口密度呈明显相关；在0.01检测水平下，有18种指标与人口密度的相关是显著的。这证实了影响人口分布的因素确实是多种多样的。

按相关系数的绝对值大小进行排序，在前5个指标中（相关系数均大于0.84），海拔指标即占了3个，经济指标占了2个，它们是作者新引入的单位国土面积国民收入和单位国土

表 1 人口分布的影响因素指标
Factors influencing population density to be analyzed in this paper

类别	序号	符号	影响因素指标	类别	序号	符号	影响因素指标
自然因素	①	A1	平地占总面积比例（%）	经济因素	⑰	TAC	总耕地面积（亩）
	②	A2	丘陵占总面积比例（%）		⑱	TC	垦殖指数（用小数表示）
	③	A3	山地占总面积比例（%）		⑲	TP	社会总产值（亿元）
	④	H	平均海拔（m）		⑳	m	人均国民收入（元/人）
	⑤	H1	海拔100m以下面积所占比例（%）		㉑	mA	单位国土面积国民收入（元/km ² ）
	⑥	H2	海拔100m以上面积所占比例（%）		㉒	FR	人均粮食产量（kg/人）
	⑦	H3	海拔300m以上面积所占比例（%）		㉓	FA	单位国土面积粮食产量（kg/km ² ）
	⑧	H4	海拔500m以上面积所占比例（%）	人口因素	㉔	B	出生率
	⑨	T	多年平均温度（℃）		㉕	D	死亡率
	⑩	P	多年平均降雨量（mm）		㉖	R	自然增长率
	⑪	W	水资源总量（10 ⁸ ×m ³ ）		㉗	S1	市镇人口比例（%）
经济因素	⑫	WA	单位国土面积水资源量（m ³ /km ² ）		㉘	TFR	总和生育率
	⑬	C1	第一产业国民生产总值所占比例（%）		㉙	y	年龄中位数
	⑭	C2	第二产业国民生产总值所占比例（%）	教育因素	㉚	WM1	文盲、半文盲占总人口的比例（%）
	⑮	C3	第三产业国民生产总值所占比例（%）		㉛	WM2	12岁以上人口中文盲、半文盲所占比例（%）
	⑯	CR	人均耕地面积（亩/人）				

1) 文中数据主要来自：《1987年全国1%人口抽样调查的主要数字》，中国统计出版社1992。《1988年中国统计年鉴》，中国统计出版社1988。《国土资源数据集》，综考会编·1987。

表 2 人口密度的单因素相关分析结果

The analysis results of correlations between
population density and its influence factors

序*	相 关 因 素	线性相 关系数	回归公式	排序	相 关 因 素	线性相 关系数	回归公式
①	海拔 100m 以下面积所占比例 (%)	0.91	$D = 99.24 + 6.08H_1$	⑰	丘陵所占总面积比例 (%)	-0.54	$D = 603.9 - 19.49A_2$
②	海拔 100m 以上面积所占比例 (%)	-0.90	$D = 680.5 - 5.83H_2$	⑱	总和生育率	-0.64	$D = 943.3 - 238.8TFR$
③	单位面积国民收入 (元/km ²)	0.88	$D = 120.5 + 4.67MA$	⑲	市籍人口比例 (%)	0.40	$D = 73.74 + 4.76S_1$
④	单位国土面积粮食产量 (kg/km ²)	0.88	$D = 49.97 + 0.002FA$	⑳	水资源总量	0.38	
⑤	海拔 300m 以上面积所占比例 (%)	-0.84	$D = 547.5 - 4.90H_3$	㉑	文盲、半文盲占总人口的比例 (%)	-0.38	
⑥	垦殖指数 (用导数计)	0.84	$D = 3.16 + 1309.8TC$	㉒	12岁以上人口中文盲、半文盲所占比例	-0.37	
⑦	年龄中位数	0.79	$D = -1831.5 + 88.1Y$	㉓	多年平均温度 (°C)	0.34	
⑧	海拔 500m 以上面积所占比例 (%)	-0.76	$D = 456.8 - 4.22H_4$	㉔	多年平均降雨 (mm)	0.23	
⑨	平地占总面积比例 (%)	0.68	$D = -32.6 + 8.51A_1$	㉕	单位面积水资源量 (m ³ /km ²)	0.11	
⑩	平均海拔 (m)	-0.68	$D = 416.02 - 0.14H$	㉖	人口死亡率	0.11	
⑪	人均国民收入 (元/人)	0.66	$D = 69.92 + 0.21m$	㉗	人均粮食产量	-0.06	
⑫	人均耕地面积 (亩/人)	-0.63	$D = 477.8 - 142.59CR$	㉘	总耕地面积	-0.05	
⑬	人口自然增长率 (%)	-0.61	$D = 1329.5 - 72.6R$	㉙	第二产业国民生产总值所占比例		
⑭	山地占总面积比例 (%)	-0.60	$D = 702.2 - 8.96A_3$	㉚	第三产业所占比例		
⑮	人口出生率 (‰)	-0.57	$D = 1454.3 - 50.4B$	㉛	社会总产值		
⑯	第一产业国民生产总值所占比例 (%)	-0.55	$D = 856.4 - 18.74C_1$				

* 序按相关系数大小排列

面积粮食产量。在前10个指标中，自然指标占了6个，其中海拔指标占了5个，另一个是平地所占国土面积的比例；经济因素指标占了3个，除了前述两单位国土经济指标外，还有垦殖指数，也属于单位国土经济指标；另外还有一个人口指标即年龄中位数，居前第7位。排在第11到第15位的分别是人均国民收入、人均耕地面积、人口自然增长率、山地占总面积的比例和人口出生率，它们与人口密度的相关系数均在0.57以上。

表 2 的相关分析结果表明，第一，自然因素、经济因素、人口因素、教育因素在0.05水平上与人口密度均呈明显相关；第二，自然因素与人口分布的相关最显著，其次是经济因素，再次是人口因素，相关最弱的是教育因素（在0.05水平上刚刚达到影响显著水平）。

1.2 自然因素

在自然因素中，对人口分布影响最大的是海拔高度，这不仅表现在在31种指标中与人口密度相关最密切的前两个指标都是海拔指标，在前10个指标中有5个是海拔指标，而且所选

用的全部 5 个海拔指标都与人口密度相关较好, 相关系数都在 0.68 以上, 最高可达 0.9 以上。影响第二大的自然因素是地形, 平地所占总面积的比例与人口密度呈正相关, 相关系数为 0.68, 山地所占总面积的比例和丘陵所占总面积的比例则与人口密度呈明显的负相关, 相关系数分别为 -0.60 和 -0.54。但多年平均温度和多年平均降雨与人口密度的相关不明显。这说明对于中国省级人口分布, 主要的自然影响因素不是温度和降雨, 而是地形地貌, 尤其是海拔。这似乎与以往的定性分析结论不一致。一般认为温度和降雨与人口分布的关系是比较紧密的, 但在中国省级人口分布中, 温度、降雨与人口密度的相关并不明显。其原因首先是中国的国土绝大部分位于温带和亚热带, 比较适于人口居住, 温度和降雨对空间尺度较大的省级人口密度影响不十分显著; 其次是其它因素如地形掩盖了温度、降雨的影响, 所分析的各省人口密度数据包含了各种因素的综合影响, 而在这种综合影响中温度、降雨的影响不很突出, 因而相关分析结果表明相关不明显。中国人口分布的主要特点是东多西少, 它是由中国从东到西海拔逐级升高而成三级台阶这一地形特点决定的; 这说明地形对中国人口分布的影响较之气候的影响更为显著。图 1 是我国 1987 年省级人口密度与海拔 100m 以下国土面积所占比例的相关图。它表明省级人口密度与海拔 100m 以下国土面积所占比例的线性关系是十分明显的。

1.3 经济因素

在经济因素中, 对省级人口密度影响最大的是单位国土经济指标, 包括单位面积国民收入, 单位国土面积粮食产量、垦殖指数, 它们与人口密度的线性相关指数分别为 0.88、0.88、0.84, 仅次于海拔指标与人口密度的相关系数; 其次是人均经济指标, 包括人均国民收入和人均耕地面积, 相关系数分别为 0.66、-0.63; 再次是第一产业国民生产总值所占比例, 相关系数为 -0.55。以上几种因素与人口密度的相关系数均较高。其它一些经济指标如人均粮食产量、总耕地面积、第二产业国民生产总值所占比例、第三产业国民生产总值所占比例、社会总产值与人口密度的相关均不明显。从产业结构与我国省级人口密度的关系来看, 第一产业与人口密度关系紧密, 第二、第三产业关系还不明显, 说明我国还是一个以第一产业为主 (近 80% 的人口为农业人口) 的国家, 第二、第三产业尚不发达, 不能吸收大量的人口, 因而与人口密度的关系也不明显。

图 2 给出了 1987 年我国省级人口密度与单位国土面积国民收入的相关图, 看出: 人口密度与单位面积国民收入呈明显的曲线相关。据此, 我们又作了二者之间的二次曲线和三次曲

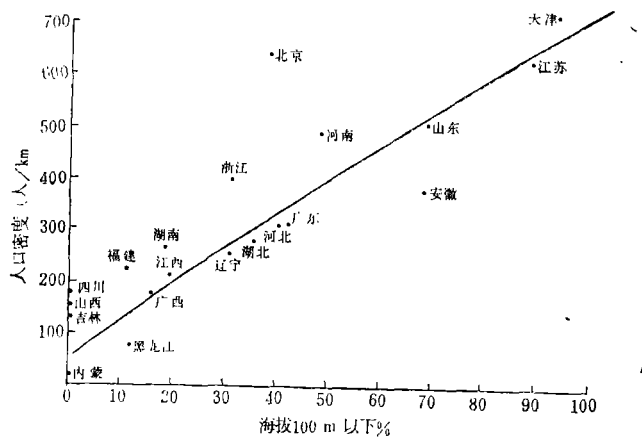


图 1 省级人口密度与海拔 100m 以下国土面积所占比例的相关图

The relationship between population density, and the percentage of territory area below the altitude of 100 metres

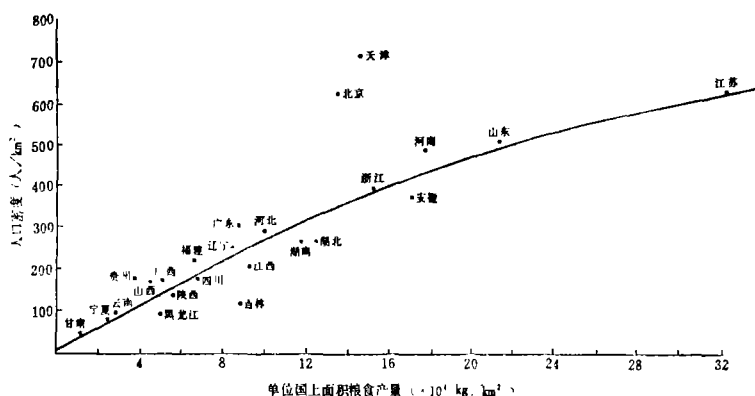


图 2 人口密度与单位面积国民收入的相关

The correlation between population density and
national income per unit territory area

线相关, 结果如下:

二次曲线: $D = 36.9 + 11.07MA - 0.0447MA^2$, $r = 0.96$

三次曲线: $D = 14.2 + 14.28MA - 0.1136MA^2 + 0.00032MA^3$, $r = 0.97$

式中 D 表示人口密度 (人/ km^2), MA 表示单位国土面积国民收入 (元/ km^2)。二次曲线和三次曲线的相关系数分别为 0.96、0.97, 比线性相关系数 0.88 要高, 说明人口密度与单位国土面积国民收入之间为曲线相关。

1.4 人口因素

在人口因素中, 除了人口死亡率 (0.11) 之外, 其它 5 个人口指标与人口密度均有明显的相关, 人口密度与人口自然增长率 (-0.61)、人口出生率 (-0.57)、总和生育率 (-0.46) 呈负相关, 与年龄中位数 (0.79)、市镇人口比例 (0.42) 呈正相关。说明人口密度越高的省份, 其年龄中位数也较高、市镇人口比例较高, 而人口自然增长率、人口出生率、总和生育率也较低。人口密度与以上人口因素的相关分析与定性分析的结论也是相符的。不过人口密度与人口因素指标的较高的相关系数, 只能说明它们之间有较好的相关关系, 而不能说它们之间存在人口因素决定人口密度的简单的因果关系。实际上, 人口密度与其它人口因素是相互影响的。例如人口密度与市镇人口比例的关系, 一方面由于市镇人口密度远大于乡村人口密度, 所以市镇人口比例高的省份的人口密度很可能会高于市镇人口比例低的省份, 另一方面人口密集的地方更容易形成市镇, 所以高的人口密度很可能产生高的市镇人口比例, 因而人口密度与市镇人口比例之间不是单向的因果关系, 而是交互影响的。

1.5 教育因素

在教育因素中, 文盲及半文盲占总人口的比例与人口密度的相关系数为 -0.38 , 12 岁以上人口中文盲及半文盲所占比例与人口密度的相关系数为 -0.37 , 在 0.05 检测水平下达到显著相关水平。与人口因素类似, 教育因素与人口密度之间也不存在单向的因果关系, 它们之间也是交互影响的, 一方面发达的教育、或者说教育发达的地区更能吸引人口, 另一方面人

口密集的地区更容易发展教育。

值得进一步探讨的是自然因素中的海拔和经济因素中的单位国土经济指标对人口密度相关最密切的真实含义。首先,我们认为,海拔100m以下国土面积所占比例与人口密度的高度正相关,不仅说明了我国人口密集于低海拔的事实,而且也说明我国人口分布仍严重地受限于地貌条件。海拔较低的平原丘陵区,条件优越,人口承载力较大;相反,海拔较高的高原和山区条件较差,人口承载力较小,限制了人口密度的提高。这说明目前的生产力水平还不能超越高海拔及山地对生产力的限制,因而不能在高海拔地区维持高密度的人口。其次,单位国土面积经济指标与人口密度的高度正相关,一方面说明了人口与经济的固有均衡性,一定的人口要以一定的经济基础为前提,因而人口密度与经济密度呈密切正相关;另一方面也恰恰反映了我国人口与经济在低水平上的均衡和互相适应。图3是1987年我国省级人口密度与单位国土面积粮食产量的相关图。显然可见,除了大城市之外,二者呈很好的正相关,只是在单位国土面积粮食产量较高时(如江苏省)才有偏离直线的趋势,而且是偏向右

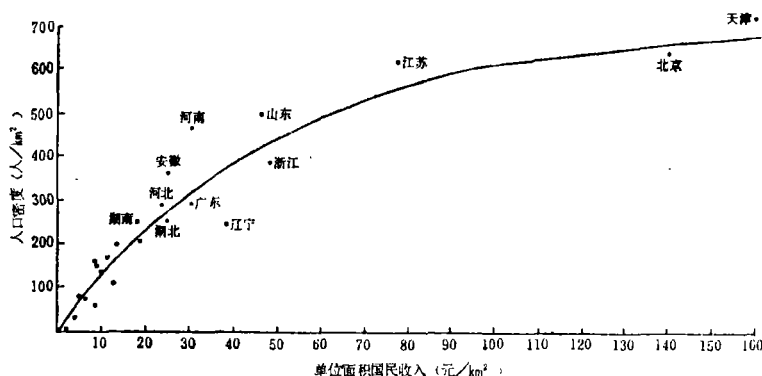


图3 人口密度与单位国土面积粮食产量相关图

The correlation between population and grain yield per unit territory area

下方。说明多数省份人口密度的高低,仍主要受到于粮食的多寡,能生产多少粮食就有多少人口,粮食生产仍停留在只能维持温饱的水平,人口对粮食有很大的压力,这构成了人口压迫生产力的又一个层面。还可看出江苏省,人口密度的增长已落后于粮食的增长,这或许说明江苏省尚有余粮以作他用,新增加的粮食已不再用于解决人们的温饱问题,或许也意味着江苏省在全国各省份中率先解除了人口对粮食的压力而进入了小康阶段。从图2也可看出我国的人口分布极大地受制于经济因素,而且具有人口压迫生产力的特点。总之,我国省级人口密度与海拔、单位国土面积粮食产量和国民收入的高度相关,既反映了人口分布对自然、经济条件的适应,又在一定程度上反映了我国人口对自然环境和社会生产力的压迫,而且还说明了我国粮食商品率很低,各省基本上都是自给自足的。

2 多因素回归分析

分析已知,有很多因素与人口密度的相关很明显。在进行多元回归时,首先应当选择影

响显著的因素,但因为这些因素之间还存在相关,为消除因素之间相关的影响,还需挑选独立的影响因素。

影响人口密度的几个主要因素之间的相关矩阵见表3。表中各符号的意义与表1相同。可以看出,与人口密度相关密切的几个主要因素,它们之间的相关也是很密切的。

表3 人口密度主要影响因素之间的相关

The coefficient between the main factor influencing population density

	HI	AI	TC	MA	FA
HI	1	0.81	0.85	0.75	0.89
AI		1	0.72	0.60	0.69
TC			1	0.62	0.87
MA				1	0.62
FA					1

经过筛选、逐步回归,得到如下公式:

$$D = 29.18 + 240.59TC + 2.785MA + 0.00106FA$$

式中: D表示人口密度(人/km²), TC为垦殖指数(以小数计), MA为单位国土面积国民收入(元/km²), FA为单位国土面积粮食产量(kg/km²)。经统计检验,在 $\alpha = 0.05$ 水平上,上式中各因素的作用是显著的。

上式的意义在于:人口密度与垦殖指数与单位国土面积国民收入、与单位国土面积粮食产量成正比。上式只显含这三项经济指标,实际上也可用其它一些指标建立表示人口密度计算公式。如果这样建立的公式反映了人口密度与其因素之间的真实联系,而且应用公式时较之建立公式时的影响人口分布的主要因素没有发生明显变化,就可以应用这样的公式计算以及预测人口分布的变化。

本文分析了31种指标与省级人口密度的相关系数,结果表明:在所分析的全部因素中,以自然因素中的海拔与人口密度的线性关系最密切;其次是经济因素,包括单位面积国民收入、单位国土面积粮食产量、垦殖指数;再次是人口因素年龄中位数;教育因素与人口密度的相关软弱,但也达到影响显著水平。单位国土面积国民收入与人口密度呈曲线相关,曲线相关系数明显高于线性相关。

人口密度与海拔、单位国土面积国民收入、单位国土面积国土产量的高度相关,一方面表明我国人口的分布主要受制于自然条件和经济条件,另一方面这种制约关系本身也表明我国人口对自然环境和社会生产力有很大的压力。

参 考 文 献

- (1) 胡焕庸、张善余. 中国人口地理. 华东师范大学出版社, 1984.

THE QUANTITATIVE ANALYSIS OF FACTORS INFLUENCING POPULATION DISTRIBUTION IN CHINA

Meng Xiangjing Jia shaofeng

(Institute of Geography, Chinese Academy of Sciences and state Planning
Commission of the People's Republic of China)

Subject terms: Quantitative Analysis, influence factor, population,
China

Abstract

The correlation between provincial population density and its 31 influencing factors are analysed. Here two new economic indexes are introduced, grain yield per unit territory area and national income per unit territory area. The results of analysis show that these two indexes have closer relationship with population density than other economic indices. within all the 31 indices analysed, natural factors, especially elevation above sea level, influence population density most prominently. Economic factors influence population density more prominently than population factors do. Educational factors have the weakest effect on population density.