

明清回族进士与回族人口空间分布¹

咸同年间的云南及西北回民事变，造成了极其严重的人口损失，探讨事变之前的回族人口规模及其空间分布状态是人口史和回族史研究的重要内容。以往对于这一问题的研究，大都基于描述性的文字史料以获得较为感性的认识，亦有基于大样本回民聚落进行的相关统计和分析。²其实，回族进士对于研究回族历史人口的规模及分布也具有重要的指标意义。

杨大业先生对明清回族文武进士进行了细致入微的历史学考证，并分门别类进行了较为系统全面的分析和总结。³但是，回族进士数据是具有时间序列特征的空间数据，单纯依靠传统历史学的研究方法，无法全面展示其空间分布状态、特性和规律。笔者曾撰文尝试引入 GIS 的手段对这一问题进行研究，取得了一些发现，但文章仍然偏向于历史学的考证与分析，缺乏对回族进士与回族人口空间分布的 GIS 展示与分析。⁴因此，有必要就此问题展开进一步的讨论。

一、明清回族进士数据特征

杨大业先生对明清回族进士的考证，按省、府、县三级行政区来逐级考证其归属地。但已考的回族进士，并不在一个统一的时间切面上，而是一个在较长时段内，逐年累积的汇总结果。每位进士的户籍、乡贯或学额属地所在的厅、县名称，也都是特定时间切面上的政区名称。研究数据的空间精度为县，时间精度为年。

进士是中国古代科举通过最后一级的考试者，对于读书人来，考取进士是一

¹ 基金项目：上海市教育发展基金会（2013 曙光项目 13SG10）；复旦大学历史地理研究中心自立项目。

² 路伟东：《同治以前陕甘回民聚落分布与数据库建设》，《西北民族研究》2012 年第 4 期。

³ 《回族研究》从 2005 年第 1 期开始，直至 2010 年第 2 期结束，期间除 2006 年第 1 期和 2009 年第 2 期外，前后共连载杨大业先生《明清回族进士考略》系列论文共 20 篇，共得明清回族文武进士凡 329 名，其中清代文武进士 244 名。

⁴ 路伟东：《GIS 支持下的小人口基数小概率历史事件研究——以清代回族进士规模与空间分布为例》，《回族研究》2014 年第 2 期。拙文在《回族研究》第 2 期发表之后，笔者感觉回族进士空间分布这一问题很多方面仍然没有表达清楚，还有进一步探讨的余地。本文就是在前文基础上的继续研究，因面对同样的数据，在特征分析和处理原则等方面有所重复，故作了适当删减。读者如想了解完整的数据处理方法，请参见前面。对此重复造成的不便，亦深表歉意。

件非常不容易的事情。有清一代，文科常科平均每省每科不足 12 名，⁵武科人数更少，每省每科仅有 4 名。⁶

相较汉民，回民人数少，分布亦很零散。长期以来，经堂教育作为回族内部主要教育形式，其教学内容基本以伊斯兰经文为主。而回民自己，又往往自视“杂居三教之间，濡染流俗，同化是惧，兢兢保守，惟恐不及。”⁷这使得回民在以儒家经典为主要内容的科举考试中，处于不利的地位。⁸因此，相对汉民，回民士子能考中进士，更属不易。

从统计学的角度讲，回族进士的产生，是小人口基数的小概率事件。对于某些州县来讲，这样的事情只要发生一、两次，就会形成一个高发生率。对于另外一些州县来讲，没有发生这种事情的结果就是零发生率。另外，因史料和考证原因造成的数据缺漏，也加剧了回族进士空间分布的不均衡性。

由于随机误差及数据缺漏的影响，研究者对这种小概率事件发生率的估计可能就不够可靠。解决这类问题，比较常用的方法一般有两种：其一、构建较大的地理单元，增加人口基数，降低研究数据的空间精度，从而人为增加数据的空间平滑度，获得较为稳定可靠的事件发生率；其二、合并多年的统计数据，降低研究数据的时间精度，人为增加数据的空间平滑度，生成一个较长时间内相对稳定的事件发生率。

以上两种方法各有优劣，笔者在数据处理过程中，综合使用了这两种方法，以期更加有效地减小因随机误差及数据缺漏等原因造成的，小人口基数小概率事件空间分布的异质性（heterogeneity）问题。

二、数据处理原则及方法

进行回族进士的空间可视化表达，我们首先需要把时间序列之内的所有数据，标准化到 1820 年这一特定的时间切面上去。⁹根据各厅县进士产生的频数高低，

⁵ 江庆柏：《清朝进士题名录》，中华书局 2007 年版，第 3 页。

⁶ 刘丹枫：《清代武进士仕途研究》，辽宁大学 2012 年未刊硕士论文，第 13 页。

⁷ 金吉堂：《中国回族史研究》，宁夏人民出版社 2000 年版，第 1 页。

⁸ 路伟东：《清代陕甘人口专题研究》，上海书店出版社 2011 年版，第 57-58 页。

⁹ 1820 年的政区是清朝极盛时期的疆域状态，空间数据相对比较完整，《大清一统志》和《中国历史地图集》皆以该年为清代标准年代。另外，复旦大学中国历史地理信息系统（CHGIS）有矢量化的 1820 年数据，可以直接使用，比较方便。

进行等级符号化。进而通过空间平滑或空间插值手段，把离散的点状数据，转换为连续的面数据，从而更直观的展示明清回族进士的空间聚合或离散趋势。

为此，我们需要对原始数据做如下调整和说明：其一、考证属地精确到县一级的，数据记录到 1820 年的县级治所点数据上。治所改名的，做相应调整。比如康熙六年文进士赛玉弦，考证归属地是山东靖海卫，¹⁰1820 年为荣成县；其二、考证属地精确到府、州一级的，数据记录到该府州的附廓县治所点数据上。比如道光二十七年文进士马先登，考证属地陕西同州府，¹¹1820 年属地调整为同州府附廓大荔县；其三、考证属地精确到省或大区一级的，数据记录到相应省级治所首府附廓县上。比如乾隆武进士沙敬业，考证属地江南，¹²1820 年属地指定为江苏江宁府上元县；其四、某府州附廓县有两个以上者，数据合并到其中一县。比如陕西长安府附廓县有长安、咸宁两县，数据统一记录到长安县。图 1 是分省汇总的标准化后的明清回族进士数据。

表 1 明清回族进士分省统计

序号	省份	县数*	进士数	序号	省份	县数*	进士数
1	陕西	12	43	11	广西	2	9
2	云南	17	41	12	浙江	2	7
3	直隶	14	40	13	奉天	2	7
4	山东	15	40	14	湖北	4	6
5	江苏	9	27	15	贵州	5	5
6	河南	13	25	16	广东	2	3
7	福建	5	25	17	四川	2	3
8	甘肃	11	21	18	湖南	2	2
9	安徽	4	12	19	江西	1	1
10	山西	3	11	20	吉林	1	1
	小计	103	285		小计	23	44
					总计	126	329

数据来源：杨大业：《明清回族进士考略（二十）》，《回族研究》2010 年第 2 期。

数据说明：*号标注所列县数为本省范围内有回族进士分布的县汇总之数；进士隶属 20 省份的具体考证及划分吴根洲教授有详细考证。¹³

¹⁰ 杨大业：《明清回族进士考略（三）》，《回族研究》2005 年第 3 期。

¹¹ 杨大业：《明清回族进士考略（八）》，《回族研究》2007 年第 1 期。

¹² 杨大业：《明清回族进士考略（十七）》，《回族研究》2009 年第 3 期。

¹³ 吴根洲：《清代进士历史地理分布研究》，《考试研究》2011 年第 3 期。

二、GIS 表达的明清回族进士空间分布及特征

标准化到 1820 年时间切面上的回族进士数据，实际上是以厅、县为单位，进行了汇总，也就是统计各厅县进士产生的总人数（频数）。有了这个标准化的数据，就可以轻松实现进士数据的等级符号化（symbolic）。等级符号化后的地图可以显示数据在空间分布上的细部节点和聚合趋势，但强调和突出的只是数据显示的一种视觉效果，对于数据本身没有任何改变。

要对数据进行修正，减少空间变异，可以使用移动搜索法进行空间平滑。在此基础上，再进行空间插值，将离散的点数据转换为连续的面数据，这样可以更好的展示和分析分县汇总的进士数据的空间分布趋势。

移动搜索法（Floating catchment area, FCA）是以某点为中心画一个圆或正方形作为滤波窗口，用窗口内的平均值（或点密度）作为该点的值。将窗口在研究区域内移动，直到得到所有位置的平均值。平均值的变动比原始观察值变动小，从而实现空间上的平滑效果。¹⁴图 1 是使用移动搜索法进行空间平滑，并用核密度估计法进行空间插值之后的回族进士分布图。

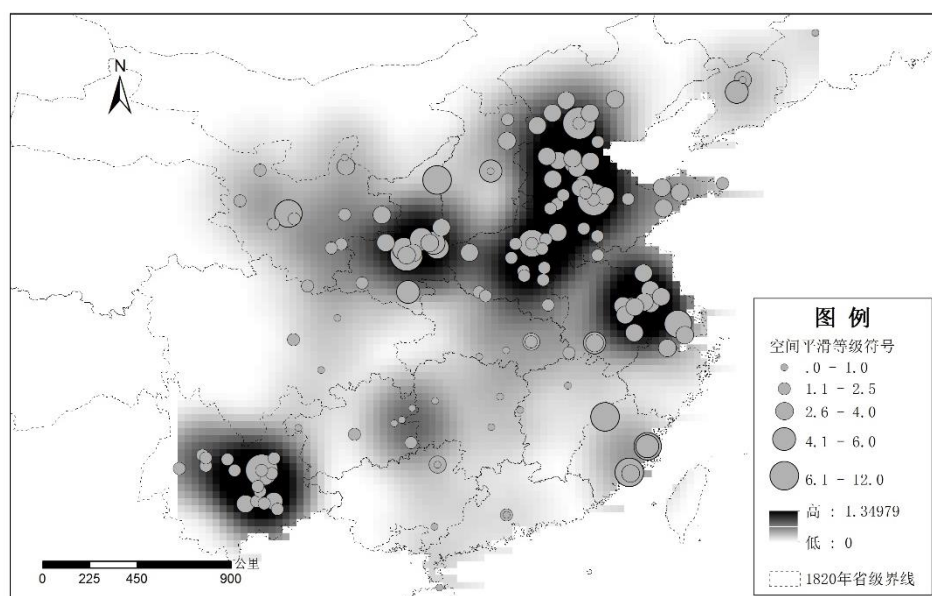


图 1 明清回族进士空间平滑与空间插值

分级符号是使用移动搜索法对数据进行空间平滑之后的效果，分级色彩则是

¹⁴ 王法辉著，姜世国、腾骏华译：《基于 GIS 的数量方法与应用》，商务印书馆 2009 年版，第 48-50 页。

使用核密度估计法对数据进行空间插值之后的效果。从图上看，分级色彩的平滑度明显要高于分级符号，这是因为，后者在前者基础上进行了数据的再平滑。从空间分布状态看，明清回族进士主要分布在长江以北地区，有大片彼此相连的聚集区。这个集中分布的区域，大体从北京往南，经直隶中南部，直达山东中西部、河南中东部地区。自此往南、往西又连接两大片区域，即往南的江苏中西部，安徽中东部片区和往西的陕西中部、甘肃中东部片区。长江以南主要分布在云南中东部、浙北杭嘉湖、湘桂黔交界处以及福建沿海地带。这其中浙北杭嘉湖实际上与又长江以北苏皖集聚区连成一片。

回族进士这一空间分布状态与我们通过传统文献获知的，清代回民人口的空间分布状态，吻合度较高，但在部分地区也有较大不同。因此，需要进行量化分析，并给予合理的历史学解释。

三、回族进士与回族人口空间分布吻合度

回族进士数据的标准化与可视化只是研究工作的第一步，接下来笔者要检验回族进士与回族人口空间分布状态之间的吻合度。或者说，经过空间平滑和空间插值处理之后的回族进士空间分布状态，在多大程度上可以反映或代表咸同以前回族人口的空间分布状态？

要回答上述问题，我们需要有同一时间切面上、同等空间尺度和精度的回族人口数据。然而，目前学界对清代回族人口数量及分布状况的了解和研究还相当有限。除了陕甘地区清代回民人口规模及变动有比较系统、深入的研究外，其他省域的相关研究成果很少。¹⁵根据既有史料和研究，我们只知道咸同以前各省回族人口的大概情况，其总数估计可能在 1000 万人左右。¹⁶这其中，仅西北地区大概就聚居了 80%左右的人口。为了进行两组数据的相关性分析，我们对回族进士按省进行汇总，具体数据见表 2。

表 2 清代分省回族人口和进士统计

¹⁵ 路伟东：《清代陕甘回民峰值人口数分析》，《回族研究》2010 年第 1 期。

¹⁶ 路伟东：《GIS 支持下的小人口基数小概率历史事件研究——以清代回族进士规模与空间分布为例》，《回族研究》2014 年第 2 期。

人口单位：万

序号*	省名	回族进士	回族人口		清回进士占比	清回人口占比
			1940年 ^①	1851年 ^②		
1	直隶	40.0	40.0	31.1	11.83%	3.01%
2	江苏	27.0	20.0	15.5	7.99%	1.50%
3	安徽	21.0	3.0	2.3	6.21%	0.23%
4	浙江	7.0	0.7	0.5	2.07%	0.05%
5	江西	1.0	0.2	0.2	0.30%	0.02%
6	福建	25.0	0.1	0.1	7.40%	0.01%
7	山东	40.0	15.0	11.7	11.83%	1.13%
8	河南	25.0	20.0	15.5	7.40%	1.50%
9	陕甘	64.0	150.0	800.0 ^③	18.93%	77.37%
10	山西	11.0	15.0	11.7	3.25%	1.13%
11	湖北	6.0	1.0	0.8	1.78%	0.08%
12	湖南	2.0	20.0	15.5	0.59%	1.50%
13	四川	3.0	15.0	11.7	0.89%	1.13%
14	广东	3.0	2.5	1.9	0.89%	0.19%
15	广西	9.0	2.0	1.6	2.66%	0.15%
16	云南	41.0	50.0	100.0 ^④	12.13%	9.67%
17	贵州	5.0	2.0	1.6	1.48%	0.15%
18	辽东	8.0	16.0	12.4	2.37%	1.20%
合计		338.0	372.5	1034.0	100.00%	100.00%

数据来源：1940年回族人口数援引自民族问题研究会编《回回民族问题》，民族出版社1982年版，第16-17页。

数据说明：*号所示省份为18个，因将八旗排除在外，陕甘合为一条记录；带圈序号说明如下：①、《回回民族问题》一书初版于1941年，¹⁷所记回族人口数据是1930年代末至1940年代初的汇总估计数，本文将其界定为1940年；②、1851年人数是在1940年人数基础上按同时期全国人口年均增长率计算而得；③、800万是研究数据，详见前所引笔者论文《清代陕甘回民峰值人口数分析》；④、100万是估计数。

对表2中的回族人口占比与回族进士占比两组数据进行线性回归分析，在 $\alpha=0.05$ 的情况下， $R=0.682$ ， $P=0.02$ 。虽然0.682的R值显示两组数据显著相关，但P值小于预设的置信度0.05，因此不接受原假设，无法对R再进一步解释。重新审视原始数据，陕甘一地人口占比最高，达到77.37%，虽然进士占比亦是最高，但也仅有18.93%。将陕甘这一异常值剔除，重新计算，结果 $R=0.592$ ， $P=0.12$ 。P值大于预设的置信度，因此接受原假设。 $R=0.592$ ，按照一般的原则，这属于显著性相关。这表明除陕甘地区外，其他地区的回族进士空间分布与回族人口空间

¹⁷ 白寿彝主编：《中国回回民族史》，中华书局2003年版，第39页。

分布之间具有较为显著的相关性。也就是说，一般情况下，回族人口多的地方，回族进士分布较为集中，反之亦然。

四、清代回族进士的空间聚集趋势

在本节笔者将分析和探讨回族进士空间的分布规律与聚集趋势，比较这种规律与趋势，是否与回族人口的空间分布规律与趋势相吻合，并给予历史学的解释，这种比较有助于我们检验前文对两者空间分布显著相关的研究结论是否正确。

为了更好的表现回族进士在空间上的分布状态和聚集趋势，笔者对空间平滑并等级符号化之后的地图做了两方面的调整：其一，在分县等级符号化图层下面叠加分省等级符号化图层。等级符号化的县级图层展现了明清回族进士空间分布的细部节点，省级图层则显示了数据在空间上的聚合趋势。需要指出的是，县与省的符号化使用了不同的等级序列，叠加分省等级符号化图层只是为了增强数据在空间分布上趋于聚合的视觉体验，并不具有真正统计的意义；其二，叠加人工运河与自然河道图层，由于缺乏道路数据，笔者暂以河道来模拟历史时期的交通网络。叠加河道图层的目的是想借此对照查看数据在不同聚集区域内是否存在各向异性（anisotropy）¹⁸，以及数据的最大离散方向是否与河道方向一致。处理之后的效果见图 2。

¹⁸ Ebdon D. *Statistics in Geography* (2nd edition with corrections). Oxford: Blackwell Publishers, 1988.

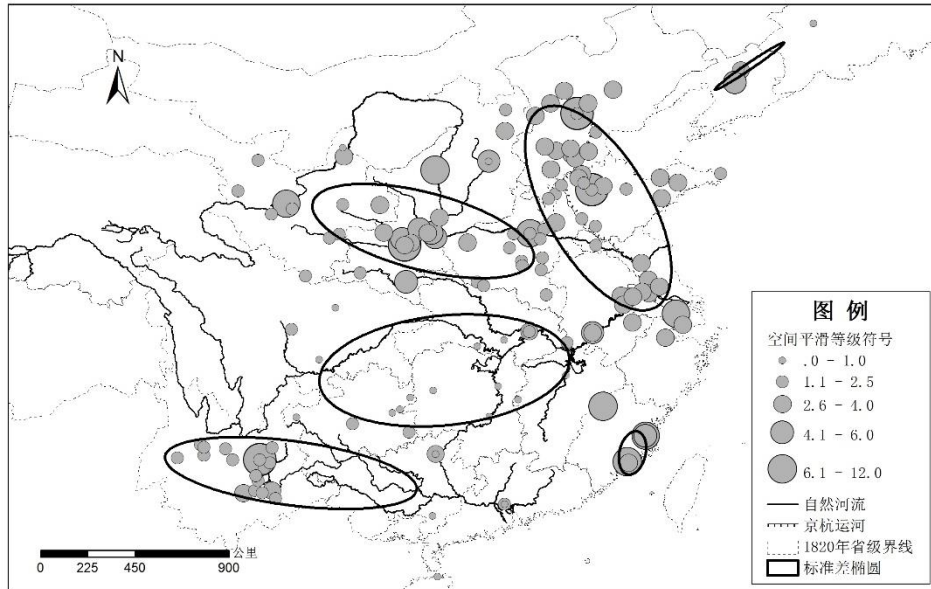


图 2 明清回族进士方向分布标准差椭圆

从图 2 看，明清回族进士空间分布比较集中的区域大概可以划分为六个区块，分别是：①、南北向的京杭运河沿线带，从北向南包括直隶、山东、安徽、江苏及浙江五省区；②、东西向的黄河中上游沿线带，从东往西包括河南、陕西、山西、甘肃等四省区；③、东西向的长江中下游沿线带，从东往西包括江西、湖北、湖南、四川、贵州五省区；④、东南沿海区，从东北往西南包括福建、广东、广西三省区；⑤、西南聚居区，包括云南一省；⑥、关外东北区，包括吉林和盛京两省区。

标准差椭圆 (Standard Deviational Ellipse) 是空间点模式分析中描述空间分布离散趋势的常用方法，¹⁹图 2 是在一个标准差下计算的县级图层标准差椭圆，黑色“10% Simple Hatch”符号化²⁰的椭圆大约包含了各自区块中 68% 的输入要素。标准差椭圆非常明确、直观的展示了数据分布的这种方向性趋向。

从图 2 看，以上 6 个回族进士相对集中的区块内部，数据在各个方向上的离散度存在明显不同，各向异性显著。与背影河道数据相对照，我们可以发现，各椭圆的长轴方向，也就是数据最大离散度方向，与人工运河、自然河道或海岸线重合度很高。这与我们通过传统文献获知的，明清以来回民人口主要集中在沿江、

¹⁹ 张志杰、彭文祥、周艺彪、庄建林、姜庆五：《空间点模式分析中离散趋势的描述研究及应用》，《中国卫生统计》2008 年第 5 期。

²⁰ ArcGIS 10 ArcMap 图层符号化样式。

沿海交通要冲之区的空间分布态势是相同的。²¹这种相同性也表明，虽然杨大业先生谦虚的声明已考回族进士只是全部回族进士的一部分，但确考的回族进士具有相当强的代表性，并且在空间分布上是比较均衡的。

五、回族进士与回族人口空间分布集中度差异度量

探讨回族进士与回族人口空间分布的集中度差异与探讨两者在空间分布上的吻合度和聚集趋势是一枚硬币的两面，分析这一问题，有助我们更加全面、深入的认识回族进士与回族人口数据在空间分布上的相互关系。在本节，笔者使用人口集中指数模型度量两者集中度差异，并通过洛伦兹曲线对这种差异进行可视化展示。主要想回答如下两个问题：1、两者的空间聚集程度是否存在差异？2、如果存在差异，这种差异有多大？

集中指数是度量某一地理要素在地域上集中程度的指标，在社会经济领域和公共卫生领域有较广泛的应用。²²地理学家使用人口集中指数(Index of population concentration, IPC)来分析人口分布集中与分散化的变动趋势。其计算公式如下：

$$\Delta P = \frac{1}{2} \sum_{i=1}^n \left| \frac{P_i}{P} - \frac{S_i}{S} \right|$$

式中 ΔP 为人口集中指数； P_i 和 S_i 为第 i 个子区域的人口数量和区域面积， P 和 S 分别为整个父区域的总人口数和总面积数。计算结果 ΔP 的值在0和1之间， ΔP 值越小，说明人口向某些地域集中的偏向越小，越趋于离散，空间分布越均匀； ΔP 值越大，说明人口向某些地域集中的程度越大，越趋于集聚，空间分布越不均匀。显然，人口集中指数是人口与面积加权的結果，较好反映了人口分布相对于土地面积的集中和分散情况。既有研究表明，人口集中指数可以作为人口空间分布特征的检测指标。²³本节研究数据见表3：

表3 分省统计回族人口、进士及面占比

序号	省名	面积占比 A	人口占比 B	进士占比 C	B-A	C-A
----	----	--------	--------	--------	-----	-----

²¹ 马金宝：《回族人口分布的地域特征简析——与其他几个少数民族的比较》，《回族研究》2000年第4期；路伟东：《清代陕甘人口专题研究》，上海书店出版社2011年版，第301-311页。

²² 魏众、B·古斯塔夫森：《中国居民医疗支出不公平性分析》，《经济研究》2005年第12期。

²³ 陈楠、林宗坚、王钦敏：《人口经济学中的GIS与定量分析方法》科学出版社2007版，第75-76页。

1	直隶	5.82%	2.91%	11.83%	0.03	0.06
2	江苏	1.55%	1.46%	7.99%	0.00	0.06
3	安徽	2.11%	0.22%	6.21%	0.02	0.04
4	浙江	1.46%	0.05%	2.07%	0.01	0.01
5	江西	2.32%	0.01%	0.30%	0.02	0.02
6	福建	2.19%	0.01%	7.40%	0.02	0.05
7	山东	2.35%	1.09%	11.83%	0.01	0.09
8	河南	2.51%	1.46%	7.40%	0.01	0.05
9	陕甘	11.27%	77.73%	18.93%	0.66	0.08
10	山西	2.91%	1.09%	3.25%	0.02	0.00
11	湖北	2.68%	0.07%	1.78%	0.03	0.01
12	湖南	3.01%	1.46%	0.59%	0.02	0.02
13	四川	8.91%	1.09%	0.89%	0.08	0.08
14	广东	3.07%	0.18%	0.89%	0.03	0.02
15	广西	3.05%	0.15%	2.66%	0.03	0.00
16	云南	5.64%	9.72%	12.13%	0.04	0.06
17	贵州	2.46%	0.15%	1.48%	0.02	0.01
18	辽东	36.67%	1.16%	2.37%	0.36	0.34
合计		100.00%	100.00%	100.00%	1.41	1.02

数据来源：各省域面积使用 CHGIS V4 数据，中国历史地理信息系统（CHGIS），复旦大学历史地理研究中心，2003 年 6 月。

计算结果显示回族人口集中指数 $\Delta P_{rk} = 0.71$ ，回族进士集中指数 $\Delta P_{js} = 0.51$ ，这表明在空间分布上，相对于回族人口，回族进士更趋于离散，空间平滑度更高一些。图 3 是用洛伦兹曲线（Lorenz curve）²⁴展示的回族人口与进士分布状态。

²⁴ 为了研究国民收入在国民之间的分配问题，美国统计学家（或说奥地利统计学家）M.O.洛伦兹（Max Otto Lorenz，1903-1907 年（或说 1905 年）提出了著名的洛伦兹曲线。意大利经济学家基尼在此基础上定义了基尼系数。

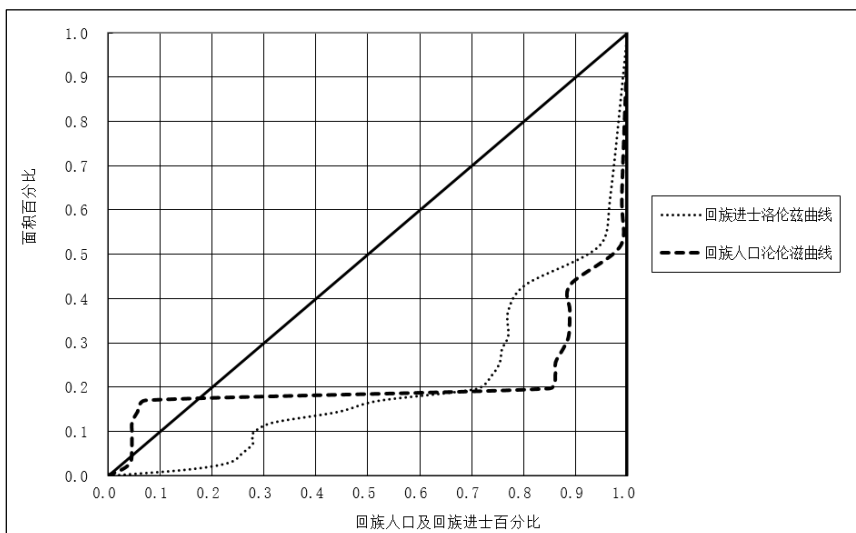


图3 回族人口与回族进士洛伦兹曲线

回族人口与回族进士洛伦兹曲线以面积排序的累计百分比为横轴，以回族人口和回族进士各自的累计百分比为纵轴。从图上3可见，超过85%的面积，只集聚了20%的回族人口。相对而言，回族进士的曲线明显要比回族人口曲线弧度更大一些，洛伦兹曲线非常直观的展现两组数据空间分布差异。

六、回族进士与回族人口空间分布不同步原因

通过上文的分析，我们可以看到，回族进士与回族人口在空间分布上的相关性与差异性就像硬币的两面，两者相辅相成，同时存在。正确理解这一问题，我们需要给予合理的历史学解释与分析。

清代的回民，是一个较为特殊的人们的群体。现有研究表明，回回人种族的或民族的认同，至少在明清之际就已经形成，就区分回、汉而言，有一条较为明确的界线。²⁵在法律层面上，回民也是一个概念非常清晰、界线非常明确的人们的群体。在立法上存在专门适用于回民的歧视性的法律条文。在具体司法实践中，回民也受到不平等的待遇。相同案件的同等罪行，对回民的处罚明显要比汉人严重的多。²⁶

²⁵ 姚大力：《“回回祖国”与回族认同的历史变迁》，《中国学术》商务印书馆2004年第1辑，第90-135页。

²⁶ 路伟东：《掌教、乡约与保甲册——清代户籍管理体系中的回民人口》，《回族研究》2010年第2期。

但是，在科举方面，回民与汉人似乎并无二致。清人称，回民“与科目、登仕版，朝廷一视同仁”，²⁷是以，回民中“有志上进者甚多，应试服官，同于士庶，而以文武科名出身，洊登显秩，为国家宣力效忠者，常不乏人”，²⁸嘉庆十九年（1814年）十二月，陕甘学政韩鼎晋以地方回民人口众多为由，奏请增设回民学额，“岁科两试，于卷面另列回民字号，凡文理稍顺，量为录取一二，以示鼓励。”²⁹此事遭到朝廷的反对，在随后的上谕中称：“回民散居各省，所在皆有，非如苗民之聚集一方者可比，湖南苗民另行编列字号，原所以示区别。若各省回民久与本处生童一体考试，如果文艺可观，原准取入学额，并无多寡限制。……所奏著毋庸议。将此谕令知之。”³⁰

由上来看，在科举考试体系中，回民既不享有苗民一样的优待政策，似乎也没有明确歧视性的限制措施。各地方回民学子与汉民学子相同，分享地方学额，靠个人能力入学晋级。在机会均等的情况下，获取名额的多少，就与人口的基数有关。理论上，人口数量较多的群体，会获得较多的份额。反之，人口数量较少的群体，获得的份额会较少。从这个角度推理，假设回民受教育水平与汉民相同，那么，就全国空间尺度来讲，回族进士在全部进士中所占的比例，大概等同于回民人口在总人口中占的比例。这是回族人口与回族进士在空间分布上存在相关性的基本逻辑前提。

但从省域空间尺度来看，决定各省区进士录取人数的核心要素不是各省的人口基数，而是学额。

学额，顾名思义，就是录取学生的额度。科举考试的目的是选拔人才，初“一切以程文为去留”，³¹即以考试优劣作为取舍的主要依据。但是，各地政治、经济、文化等诸多方面的发展水平存在差异，人口、族群及财税等项亦不相同。当局者在面对统辖区域内情形各殊的行政区域时，保证各地区都享有适度话语权与参与度的程序平等，往往可能比选拔人才本身更重要。于是，在什么样的空间尺度里选拔所需要的人才，或者说，有限的录取名额如何分配到帝国的各个区域里，就

²⁷ 杨毓秀：《平回志》，上海史学会主编《回民起义》（中国近代史料丛刊）第3册，世纪出版集团（上海人民出版社、上海书店出版社）2000年版，第59页。

²⁸ 《清世宗实录》卷九十四，“雍正八年甲戌”条。

²⁹ 韩鼎晋：《敬陈陕西情形折》，中国第一历史档案馆，档案号03-2500-041

³⁰ 中国第一历史档案馆编：《嘉庆道光两朝上谕档》，广西师范大学出版社2000年版，第19册，第981页。

³¹ 陆游：《老学庵笔记》卷五。

成了科举考试中必须要面对的制度层面的问题。

清代科举，分省取士是一项基本政策，学额的分配方式，一般按省、府、州、县行政等级逐层分拨细化。考试共分童试³²、乡试、会试三级，其中前两级考试录取考生都有额定人数。清人所说的学额通常指的是童试录取名额，³³一般大府二十名，大州县十五名，中州县十二名，小州县七、八名。咸同年间，又允许捐资增加名额。³⁴

第三级考试会试取中人数虽没有明确的额度，个别省区不同年份间的录取人数可能差额还相当大，但会试中额人数是在有明确学额的童试、乡试两级考试的基础上产生的。因此，实际操作中，大部分年份，各省每届取中的进士数量相对比较稳定。³⁵各省进士人数的多少，大体反映了各省学额分配的比例。

现有的研究表明，钱粮、丁口是划分学额的两类基本依据。这其中，钱粮作为衡量一个地方社会经济发展最综合的指标，是确定学额等级最原始、最深层的依据。³⁶钱粮、丁口，通俗的讲就是纳税额度和人口基数，也就是综合考量地方经济发展水平和人口规模，显然，这是一个综合指标。这一指标，一方面可以使那些承担较多赋税，人口密集的经济发达地区得到相对较多的份额。另一方面，也可以使那些发展落后、人口稀少的边远地区分享一定的代表权和参与度。这在一定程度上，解决了学额在不同地区之间的平衡问题。

由上来看，学额的分配虽然考虑了人口的因素，但又和人口没有直接关系。人口多的省份获得的入学额度不一定就比人口少的省份多。继续推理，回民人口多的省份，回族进士产生的数量，也不一定比那些回族人口相对较少的省份多。正是这种原因，导致了回族进士在空间分布上，比回族人口更加平滑。

³² “童试”又称“童生试”，俗称小试，是明清两代取得生员资格的入学考试，是读书士子的进身之始，应试者不论年龄大小统称童生。童试实际包含三级考试，分别是知县主持的县试、知府主持的府试和学政主持的院试。县、府两级的考试，没有名额限制，但县试通过者才有资格参加府试，府试又通过者，才有资格参加院试。院试通过者称“生员”，也就是俗称的秀才。

³³ 清代官书中所指学额相当复杂，其中包括府、州、县、厅等官学中四类生员的名额：一、廪膳生员名额，俗称廪生；二、增广生员名额，俗称增生；三、附学生员名额，俗称附生；四、贡生名额。每次考试，先录取附生，由附生递补增生，增生递补廪生，廪生依次贡入国子监，称贡生。廪生、增生名额相同，而且人数固定。惟附生是学政每来一次，就录取一次，他们的人数较多，是生员的主要部分。

³⁴ 徐根友：《武举制度史略》，苏州大学出版社 1997 年版，第 62 页。

³⁵ 谢海涛：《科举录取名额分配制度发展形成的历史》，《福建论坛·人文社会科学版》2009 年第 1 期。

³⁶ 梁志平、张伟然：《清代府州县学学额及专设学额的运作：基于长三角地区的研究》，《中国历史地理理论丛》，2011 年第 1 期。

七、结论

本文使用合并多年统计数据,生成一个较长时间段内相对稳定事件发生率的方法,对回族进士这种小人口基数小概率历史事件发生的频率进行了处理。并在此基础上,使用 GIS 手段,对回族进士与回族人口的空间分布进行了比较研究。通过分析两者的相关性与差异性,主要想探讨已考回族进士数据对于研究历史时期回族人口的规模及空间分布是否具有指标意义。

研究显示,两者的空间分布吻合度较高,相关性较为显著。除个别地区外,回族进士的空间分布可以大体反映回族人口的空间分布状态。同是,两者差异性也较明显,相对来讲,回族人口的空间集中度更高一些,而回族进士在空间更趋于平滑。造成这一现象的主要原因是清代分省取士的学额分配制度。

研究表明,目前已考的回族进士具有较强的代表性,相信随着研究的深入,可考回族进士的数量会不断增加,回族进士数据对于研究历史时期回族人口规模及分布的指标意义也会更显著。