

摘要

本文的目的是分析各种语言的时间和地理分布，并预测未来50年的发展和分布，为公司办公室的选址决策提供理论依据和建议。

首先，我们分析了语言用户的定量时间分布。考虑了影响语言用户数量的各种影响因素，并将其总结为十项指标，例如人均GDP，平均受教育年限。通过主成分分析，它们被分为四个主要成分：经济发展水平，社会平等水平，国民福利水平和文化交流。在此基础上，为母语者和非母语者建立了短期差异模型。一阶自回归模型（AR（1））用于拟合母语者的时间分布，以反映自相关特征。大多数母语使用者与非平稳单位的根过程一致。然后，我们在主要成分和第二语言用户之间建立了协整关系。建立了纠错模型，发现随机误差和纠错项均达到了稳定性。在协整空间中，主分量对L2的影响具有一阶微分平稳性质。

基于短期模式I，我们进一步建立了长期差分模型。将母语使用者的变化过程视为类似于自然人口增长的逻辑模型，该系统对于正常范围内的系数是稳定的，并且在当前条件下给定的稳定均衡具有最大容量。时间分布在很大程度上与人口的自然变化过程保持同步。此外，由于各种因素的线性组合的差异平滑性，长期以来对第二语言用户的影响被视为常数。因此，L2的模型是一个常数系数微分方程，其时间路径由语言影响的强度决定。因此，我们将L1和L2相加，以计算语言使用者的总数，这是一个非平稳动态系统。这意味着特定语言的驱动力是母语使用者的内生性增长和作为第二语言的外部影响。语言总数的时间分布正在增加。

其次，我们使用长期模型来预测每个国家未来50年的情况。有影响力的语言使用者数量显著增加，而增长方式则受第二语言传播的驱动。不太有影响力的语言使用者的数量增长不明显甚至减少。增长方式是母语人士所固有的。在前10名中，母语人士将来的数量可能会显著下降，或者非母语人士的数量可能增长不足，因此未来的排名可能会被取代。敏感性分析和蒙特卡洛鲁棒性仿真表明，我们的模型是鲁棒的和可预测的。

第三，我们建立了一个马尔可夫模型来分析语言的地理分布及其变化。本文构建了移民的过渡矩阵。根据人口增长，自然增长率和语言分布的信息，推断每个国家每种语言总数的分布。然后，我们集中精力查看首都的形象。预测表明，语言的地理分布在将来往往会交织在一起并以第二语言传播。

接下来，我们通过聚类分析找到新的办事处，并认为需要具备说英语和中文的能力，并且还考虑了经济的发展。因此，我们构建了四个指标，说英语和说中文的比率，人均GDP和净移民。短期和长期模型以及预测结果的马尔可夫模型计算值。根据这四个指标，我们分别通过聚类分析来分析这224个国家。由于每个国家等级的量化结果，我们使用模糊评估的多目标决策（MODM）来计算等级，并选择6个等级最高的国家首都作为新办事处的所在地。

最后，使用MINE模型对经纬度坐标网格进行网格划分。然后计算网格密度指数，以识别办公室分布密集的区域。消除适当的办公室数量可以减少服务成本，因为服务全球规模最大，办公室位置最少。结果表明，设立四个新办事处更为合适。因此，就短期而言，我们建议在伦敦，新加坡，渥太华和堪培拉建立4个新办事处。从长远来看，推荐新加坡，伦敦，堪培拉和巴黎。

关键字：自回归模型，协整，微分系统，马尔可夫链，聚类分析，MINE模型

