

叶昕教授课题组代表性成果介绍

出租车乘客目的地选择模型研究

研究背景

目的地选择模型从个体出行特性出发，能够同时吸纳和权衡多个阻抗因素和小区吸引因素，用非集计思路重新考虑传统“四阶段”模型中的出行分布阶段。

模型理论

目的地选择的效用方程

目的地小区选项一般采用如下随机效用方程：

$$U_{ij} = V_{ij} + \varepsilon_{ij} = t_{ij}\beta_1 + x_j\beta_2 + \ln(z_j\delta) + \varepsilon_{ij}$$

这里的*i*代表出行者的起点小区，*j*代表出行者的目的地小区；

V_{ij} ：从小区*i*出发，选择小区*j*为出行目的地的系统效用部分；

t_{ij} ：小区*i*到小区*j*之间的出行阻抗向量；

x_j ：目的地小区*j*的非规模吸引向量；

z_j ：目的地小区*j*的规模吸引向量；

β_1, β_2 和 δ ：变量相对应的系数向量；

ε_{ij} ：相互独立且服从标准Gumbel分布的随机效用部分。

目的地选择的概率

根据分布假设和效用最大化原理可以导出基于多元logit模型的目的地选择的概率表达式：

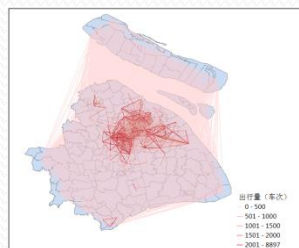
$$P_{ik} = \frac{\exp(V_{ik})}{\sum_{j=1}^J \exp(V_{ij})}$$

这里的 P_{ik} 代表从*i*小区出发的出行者选择*k*小区作为出行目的地的概率。

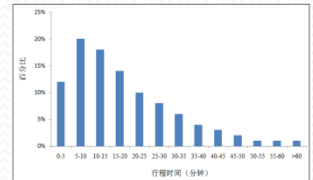
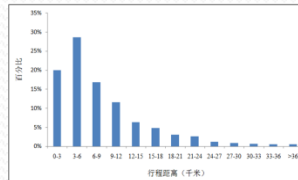
案例分析

样本构成

采用2015年4月上海市出租车GPS数据，原始数据中包含35000多辆出租车，占上海市出租车总数的70%以上，采样频率为10-30秒一次。



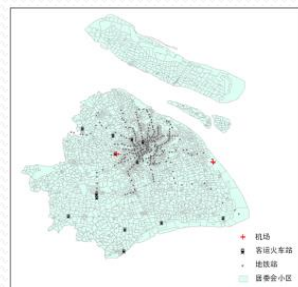
从早高峰7:00-9:00载客行程OD记录中随机抽取5000条作为PA出行样本，用于模型估计，长度分布如下：



解释变量

该模型所考虑的变量含义及数据来源如下：

特性变量类型	变量名称	含义	数据来源
阻抗变量	time	行程时间 (单位：分钟)	TransCAD软件生成的小区间阻抗矩阵和上海市出租车计价标准
	cost	花费 (单位：元)	
目的地小区非规模吸引变量	airport railway metro	机场，客运火车站，地铁站	小区内包含，则取1，否则取0
目的地小区规模吸引变量	population employment	人口数，工作岗位数	上海市第六次人口普查数据和手机信令数据



模型估计结果

应用NLOGIT5.0软件估计模型系数。

编号	变量名称	系数	标准差	T检验值
1	time (min)	-0.0882	0.0017	-51.88
2	cost (Yuan)	-0.0294	0.0012	-24.50
3	airport	4.8884	0.2228	21.94
4	railway	0.6788	0.1967	3.45
5	metro	0.4063	0.0547	7.43
“ln () ”内的规模变量				
6	population	1.0000	--	--
7	employment	4.9792	0.3978	12.52

样本量：5000

Log-likelihood: -6087.960

Rho-square (ρ^2): 0.5931

Adjusted rho-square (ρ^2): 0.5927

变量系数讨论

行程时间和花费前的系数为负，表明行程时间越长、花费越多，出行阻抗越大。

非规模吸引变量前的系数均为正，表明小区内包含的机场、客运火车站、地铁站会增加该小区对于出租车乘客的吸引力。

与小区人口数相比，工作岗位数对于出租车乘客的目的地选择产生更强的吸引。