

博客流 平均误差6000人

星期数等因素,数据来源于世博局前一个月或前几天的数据统计

$$\epsilon_t \sim W(0, \sigma^2)$$

8月预测和实际数据比较

日期	8.1	8.2	8.3	8.4	8.5	8.6	8.7	8.8	8.9	8.10	8.11	8.12	8.13
预测(万人)	39.89	33.86	34.26	35.69	35.77	38.81	43.96	39.02	40.65	41.05	36.50	30.96	34.53
实际(万人)	31.6	33.67	33.6	33.57	35.21	38.81	44.24	39.07	39.84	42.27	37.38	36.97	



张惠楚预测 7月31日客流实例

根据时间序列分析模型计算,以最高温度为温度的标志,首先研究数据发现在温度上升到一定程度后,每上升1℃,人数会急剧下降,于是很自然地与二次曲线的性质结合在一起。如抽取了最低点、最高点和中间点三个有代表性的数据。29℃——45万人,34℃——41万人,39℃——32万人,取7月份的31天人数的平均值为一标准量,为44.63万,建立方程,用Matlab解此三元一次方程。

此外,星期数对客流也有影响,如星期六根据计算就要用上15%的人数加成,预测得到的7月31日的值为48.86万,因为7月31日为星期六所以将乘上调整系数115%,然后当天最高气温为37.4℃,再根据先前求得温度函数进行调整,得到最终值为44.61万,实际数值为44.09万,误差为0.52万。

AR模型的数学定义

输出变量序列 $\{y_t\}$ 是它的自回归和移动平均的线性函数,其表达式为:

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + \epsilon_t \quad (1)$$

【附加说明】自回归模型(AR)式(1)为自回归模型,其中 $\alpha_1, \alpha_2, \dots, \alpha_p$ 为自回归系数,是模型的特征参数,而 ϵ_t 是相互独立同分布序列,且服从均值为0,方差为 σ^2 的正态分布。自回归模型(AR)式(1)中 $\alpha_1, \alpha_2, \dots, \alpha_p$ 不相等,不是一般性。在(1)中假定序列 $\{y_t\}$ 均值为0,若 $\mu \neq 0$,则令 $y_t = y_t - \mu$,可得(1)式的形式。记 $\alpha_1, \alpha_2, \dots, \alpha_p$ 为自回归系数,即 $\alpha_1, \alpha_2, \dots, \alpha_p$ 可表示为:

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + \epsilon_t \quad (2)$$

令 $\alpha_0 = 1 - \alpha_1 - \alpha_2 - \dots - \alpha_p$

模型可写为:

$$\alpha(B)y_t = \epsilon_t \quad (3)$$

根(1)应满足平稳性条件多项式 $\alpha(B)$ 的根均在单位圆外,即 $|\alpha_i| < 1$ 的根式(1)。

▲张惠楚的论文节选。

◀8月5日,上海世博园白莲泾入口,入园的游客奔向热门场馆。

早报记者 王炬亮 图

世博局:钻研精神要肯定 但不鼓励市民完全借鉴预测数据

“学生从做调研和写论文的角度而言无可厚非,但是作为市民的出行参考依据还是值得商榷的,有可能会造成误读。”

“尽管预测不能做到百分之百准确,但是作为一名中学生,能够在这么短的时间里做出这样的成绩,是很不错的。”

早报记者 韩晓蓉 实习生 杨鑫健

昨天,上海世博局运营指挥中心副总指挥张清在听了早报记者的介绍后,对张惠楚积极钻研的精神也给予了肯定,并表示,张惠楚的计算预测的相关参考因素的确和世博局的专业人员的计算模型之间有一定的重合,不过,由于张惠楚掌握的数据比较少,其预测的数据并不鼓励市民完全借鉴。

据了解,世博局有一个包括了市

统计局专家在内的近三十人的客流预测的专家队伍,数据来源比较全面,除了气温、星期数等基本情况外,还有市里提供的宾馆旅馆的住宿情况,上海各主要道口的出入情况等权威的数据,世博局的客流专家预测团队每天下午5点左右会将相关数据进行整理分析,并就初步出炉的数据进行研判。

张清提醒,学生从做调研和写论文的角度而言无可厚非,但是作为市民的出行参考依据还是值得商榷的,有可能会造成误读。

张清表示,客流预测达到相关重大节点时会有一些公布,比如前一段时间出现的50万以上的大客流,世博局就及时通过媒体发布,告之游客,但是否每天的预测客流都需要提前一天预告,目前方案还不成熟。

对此,张惠楚坦然地表示,自己在用这个方式进行计算时,只考虑了气温、星期数等极少的数据,而且只用了一个月的官方数据值作为支撑,因此并不可能完全正确,自己只是把和教授及研究生们一起做科研,一起

思索课题作为自己提前体验大学的一种快乐体验,并不看重自己的预测会带来什么样的实际参考作用。

上海世博局运营指挥中心副总指挥张清也坦言,客流的预测做到百分之百准确是不太现实的,世博局的客流预测也存在一定的误差,少的时候误差一两千人,多的时候误差一万人以上都是存在的。世博局内部的客流预测信息除了前一天晚上5点的数据外,在第二天的上午11点还会做一些修正。