

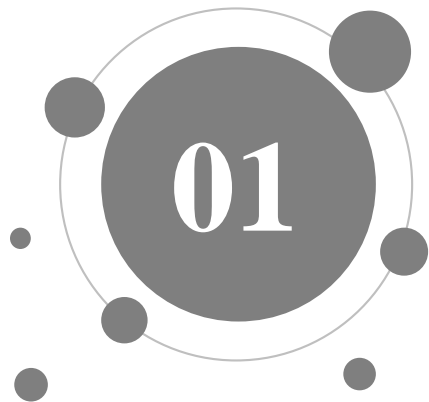
全国新型冠状病毒(COVID-19) 疫情预测分析

狗熊会 | 精品案例

主讲人：灰灰



- 01.背景介绍
- 02.数据说明与描述
- 03.增长率分析
- 04.模型分析
- 05.结论与展望



PART ONE

背景介绍



背景介绍

武汉疫情

- 国内新冠疫情爆发于湖北省武汉市
- 武汉1月23日封城，共计76天
- 截止至5月13日24时：
 - ✓ 武汉市累计确诊50,339例，占全国59.60%
 - ✓ 湖北省累计确诊68,134例，占全国80.66%

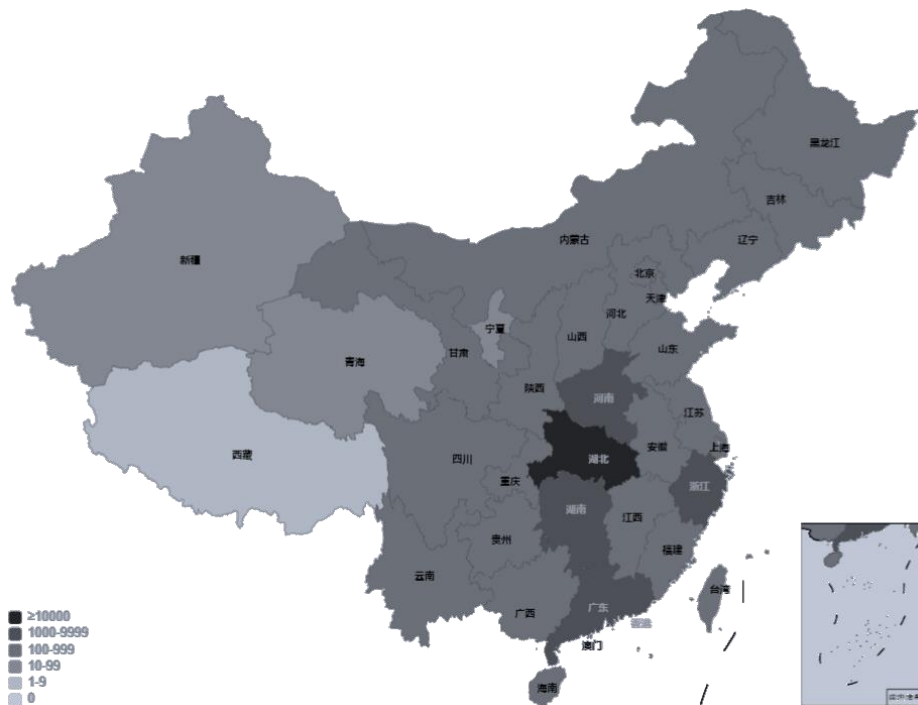




背景介绍

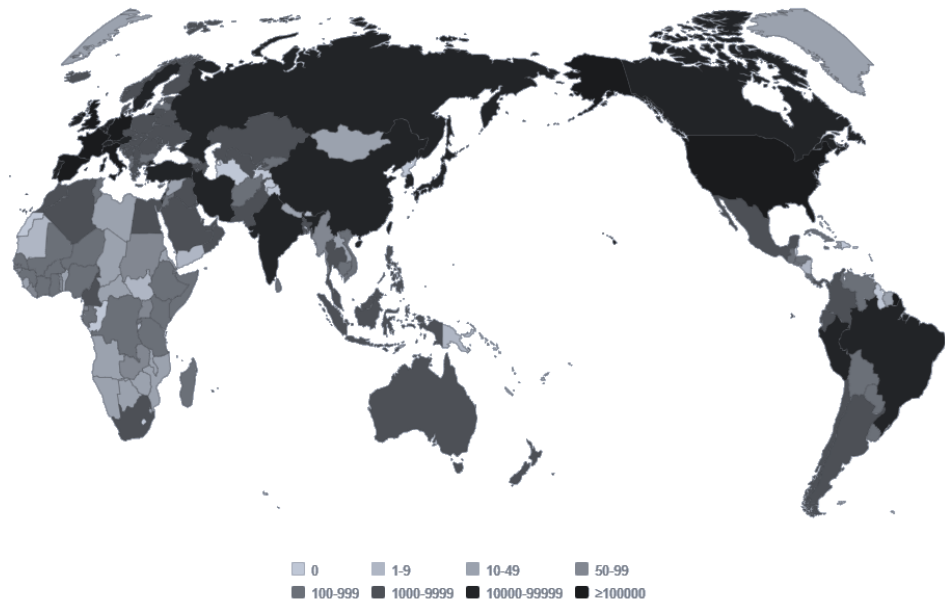
国内疫情

- 2020年1月23日始，除西藏外，所有省份均启动“重大突发公共卫生事件I级响应”
- 截止至5月13日24时：
 - ✓ 全国现有确诊199例，占全球0.01%
 - ✓ 全国累计确诊84,464例，占全球3.71%
 - ✓ 全国累计死亡4,644例，占全球1.56%



国外疫情

- 截止至5月13日24时：
 - ✓ 国外累计确诊4,342,252例，现有确诊2,472,411例，累计死亡293,506例
 - ✓ 美国累计确诊数量最多，占全球31.30%
 - ✓ 法国病死率最高，为19.57%
 - ✓ 13个国家单日新增确诊病例破千，美国单日新增逾2万例，俄罗斯单日新增逾1万例
 - ✓ 印度疫情发展迅猛，3天突破6万例





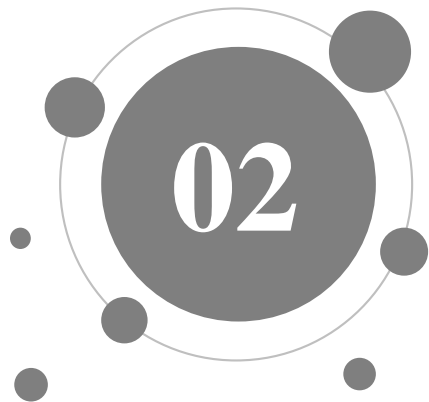
背景介绍

研究问题

- **每日新增确诊**人数指标的走势直接反映了疫情的发展态势，因此希望对该指标进行精确

研究意义

- 决策者：更好地了解全国疫情发展规律，并据此做出科学的决策，如提供合理的复工指导、适时调整响应机制
- 医学工作者：更好地把握疫情的发展变化规律，从而实施更加及时有效的医学举措
- 普通民众：建立合理预期，对自身心理状态进行及时疏导与调适，从而采取更加积极的应对方式



PART TWO

数据说明与描述

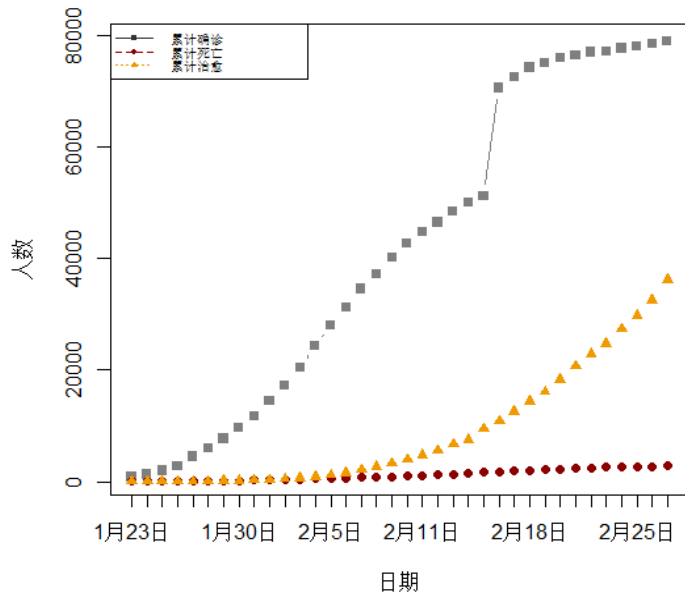
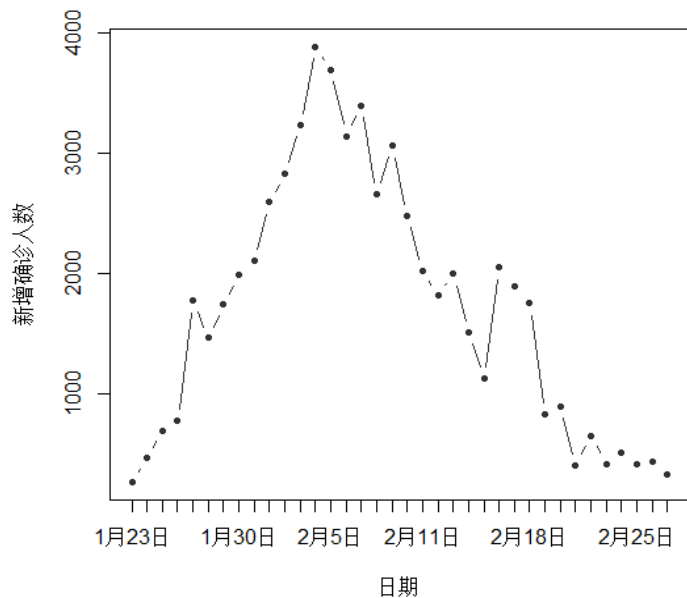
原始数据说明

- 共8个变量，原始数据来自国家卫健委官网
- 记录2020年1月23日-2020年2月27日期间全国每日疫情情况

字段名	字段说明	取值范围
新增确诊人数	t日新增确诊的人数	77 ~ 3,887
累计确诊人数	t日累计确诊的人数	291 ~ 78,824
现有重症人数	t日现有重症的人数	95 ~ 11,977
累计死亡人数	t日累计死亡的人数	6 ~ 2,788
累计治愈人数	t日累计治愈的人数	34 ~ 36,117
累计密切接触人数	t日累计密切接触的人数	9,507 ~ 656,054
正在医学观察人数	t日正在接受医学观察的人数	8,420 ~ 189,660
新增密切接触人数	t日新增的密切接触的人数	3,610 ~ 31,470

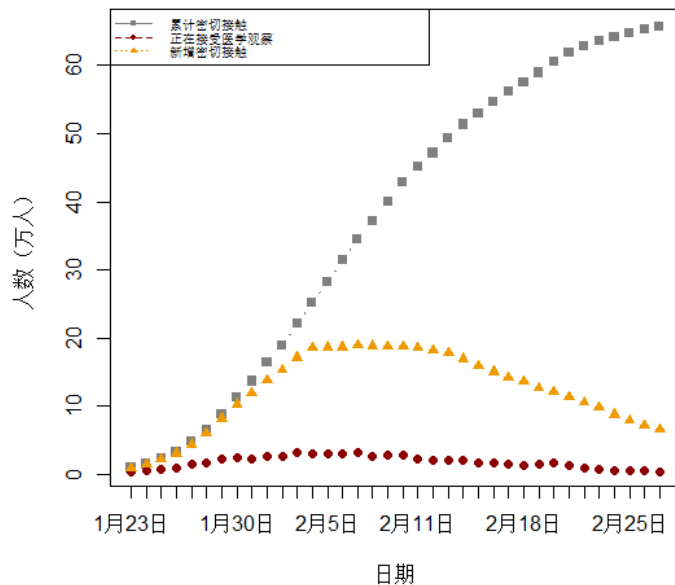
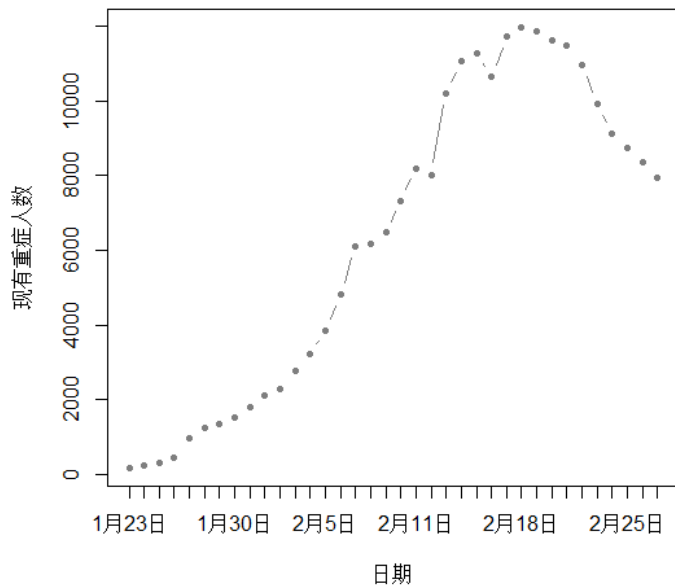
原始数据描述

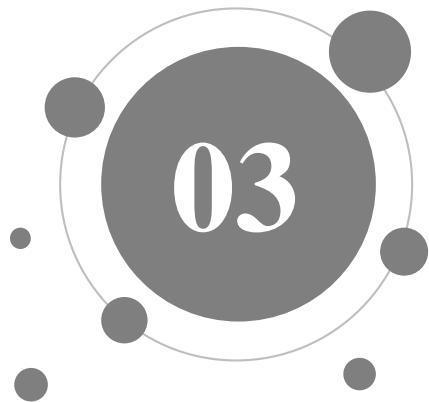
- 新增确诊人数急速上升后急速下降
- 累积确诊人数逐渐平稳，累积治愈人数迅速上升



原始数据描述

- 2月18日重症人数达到顶峰，为11,977例
- 累计人数逐渐攀升，正在医学观察的人数出现波峰后缓慢下降，新增人数变化较为平稳





PART THREE

增长率分析



增长率分析

新增确诊增长率

- 新增确诊病例数可以衡量疫情发展的严重程度，新增确诊病例的**增长率**则是指示疫情发展状态的风向标，对日后疫情发展预测有重要意义
- 研究中将“ $t+1$ 日新增确诊增长率”作为因变量。基于原始数据，其具体定义如下：

$$t+1\text{日新增确诊增长率} = \frac{t+1\text{日新增确诊人数} - t\text{日新增确诊人数}}{t\text{日新增确诊人数}}$$



增长率分析

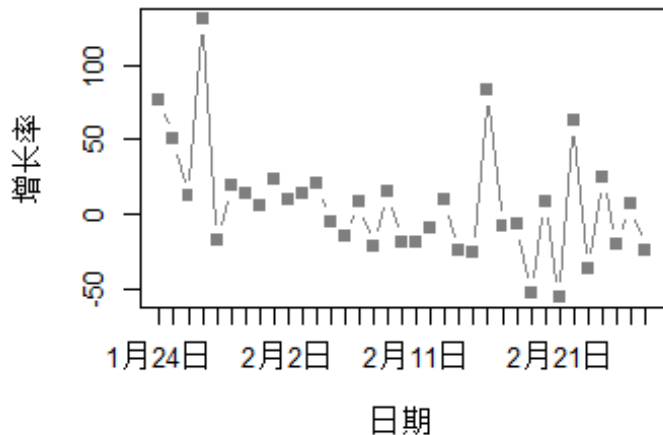
构建其他增长率指标

变量类型	变量名	说明	取值范围
因变量	t+1日新增确诊增长率	t+1日的新增确诊人数相对今天的增长率	-0.553 ~ 1.303
自变量	新增确诊增长率	t日的新增确诊人数相对昨天的增长率	-0.553 ~ 1.303
	累计确诊增长率	t日的累计确诊人数相对昨天的增长率	0.003 ~ 0.645
	现有重症增长率	t日的现有重症人数相对昨天的增长率	-0.096 ~ 1.117
	累计死亡增长率	t日的累计死亡人数相对昨天的增长率	0.011 ~ 0.640
	累计治愈增长率	t日的累计治愈人数相对昨天的增长率	0.041 ~ 0.717
	累计密切接触增长率	t日的累计密切接触人数相对昨天的增长率	-0.006 ~ 0.599
	正在医学观察增长率	t日的累计医学观察人数相对昨天的增长率	-0.100 ~ 0.659
	新增密切接触增长率	t日的新增密切接触者人数相对昨天的增长率	-0.270 ~ 0.605

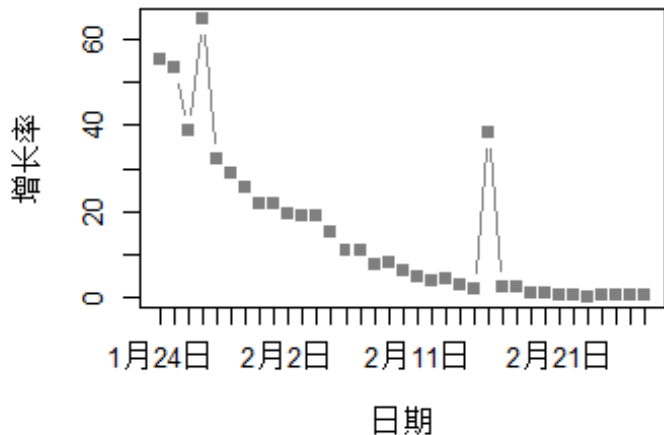
各增长率指标描述性分析

- 新增确诊增长率整体波动较大，1月23日至1月27日期间变化剧烈，至2月16日走势平稳，后至2月27日变化加剧
- 累计确诊增长率在2月16日数据激增，此后缓步下降

每日新增确诊增长率



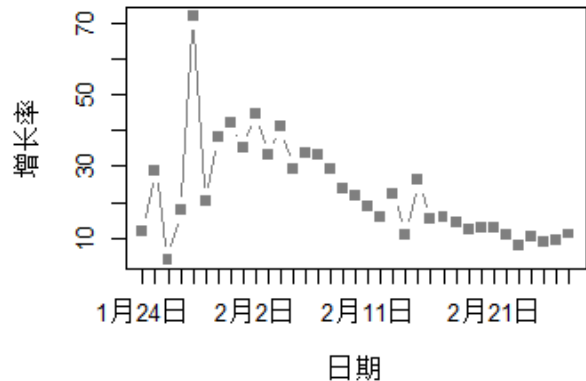
每日累计确诊增长率



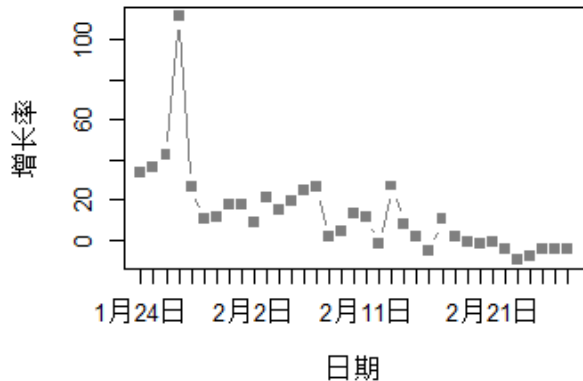
各增长率指标描述性分析

- 累计治愈增长率、现有重症增长率、新增密切接触增长率的走势呈现震荡下行趋势

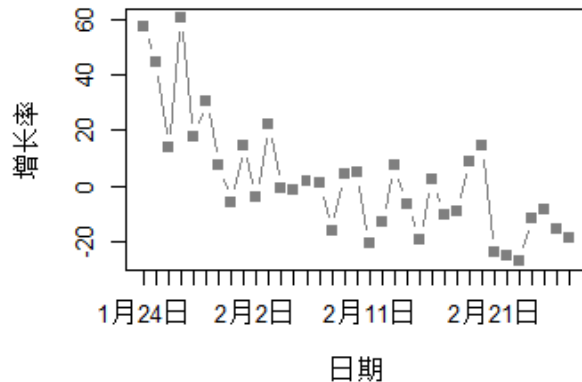
每日累计治愈增长率



每日现有重症增长率



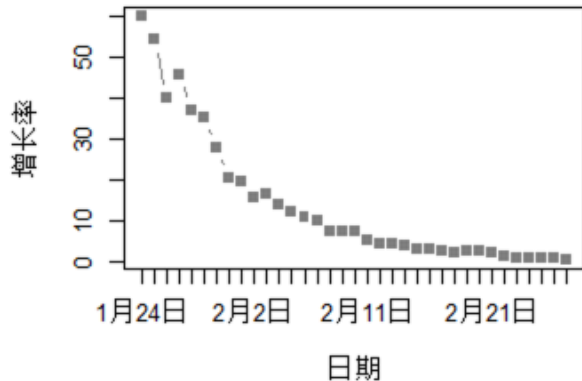
每日新增密切接触增长率



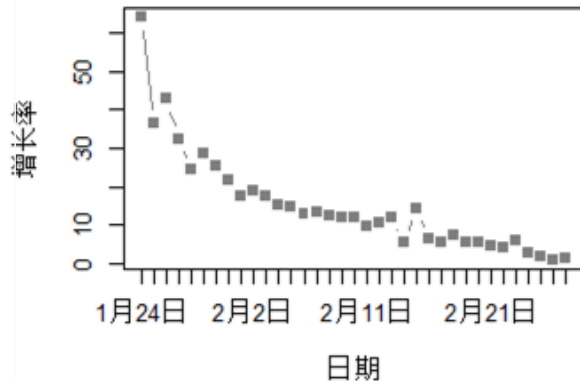
各增长率指标描述性分析

- 累计密切接触增长率、累计死亡增长率、正在接受医学观察增长率下降态势明显，后期逐渐平稳，说明疫情得到有效控制

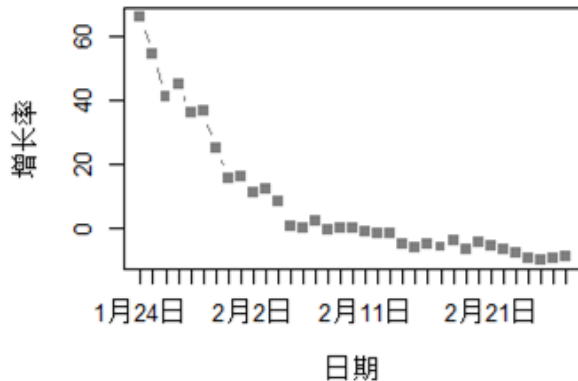
每日累计密切接触增长率



每日累计死亡增长率

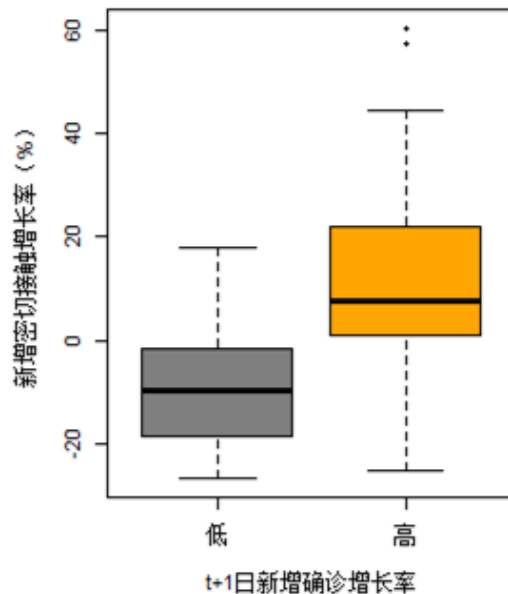
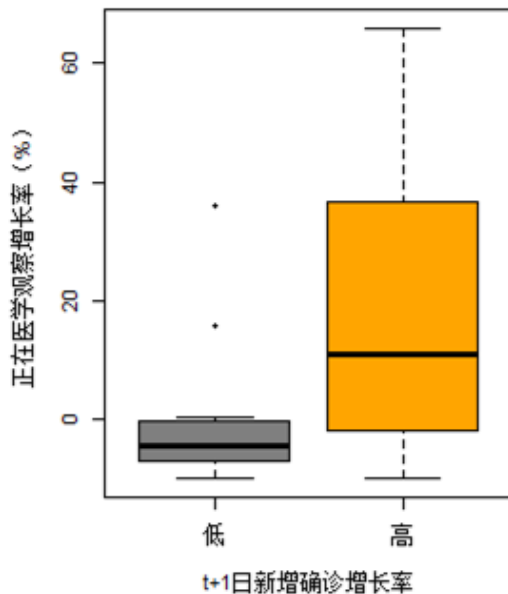
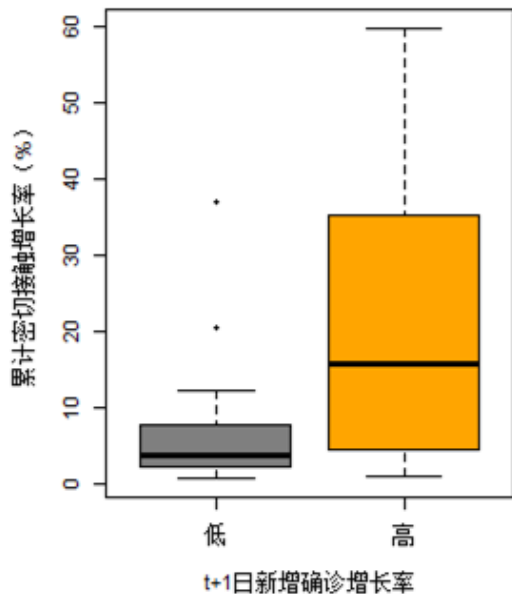


每日正在接受医学观察增长率



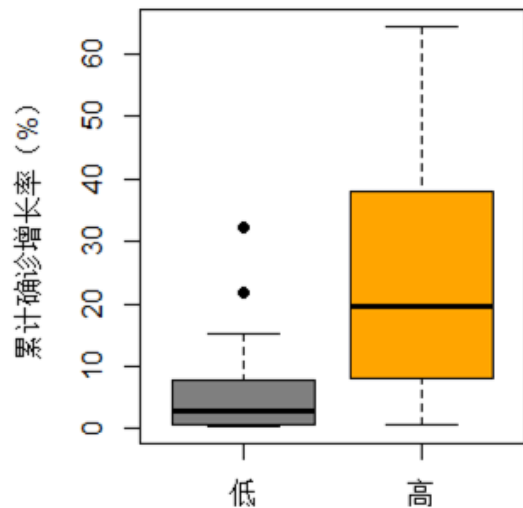
各指标与因变量的关系

- 累积密切接触增长率、正在医学观察增长率和新增密切接触增长率在新增确诊增长率的高水平组中取值都相对较高

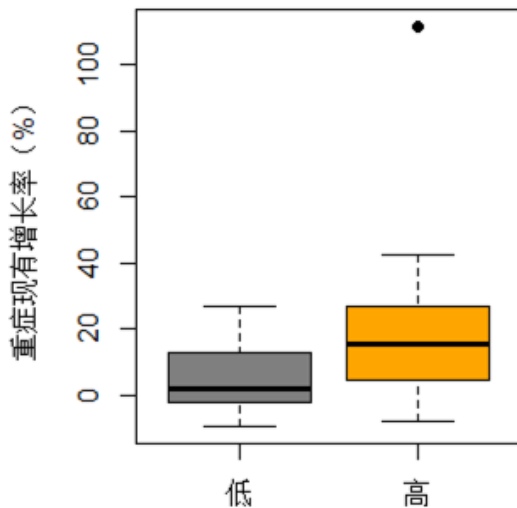


各指标与因变量的关系

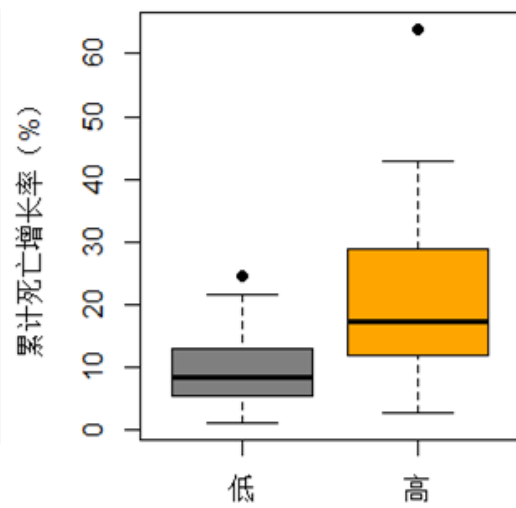
- 累计确诊增长率在因变量高低两组间分布差异较明显，与新增确诊增长率保持较为一致的变化趋势
- 现有重症增长率、累积死亡增长率在因变量高水平组中相应的取值也较高



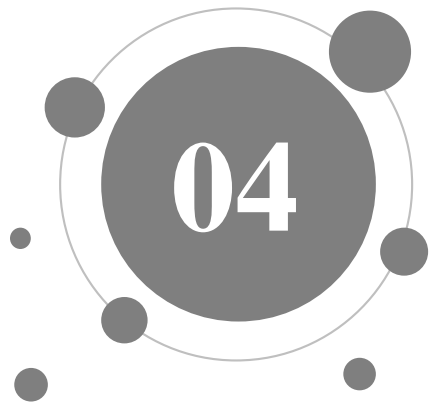
t+1日新增确诊增长率



t+1日新增确诊增长率



t+1日新增确诊增长率



PART FOUR

模型分析

模型结果

- 以 $t+1$ 日的新增确诊增长率为因变量，以其他增长率变量的 t 日取值为自变量，建立线性回归模型，并通过AIC准则进行变量选择

变量名称	AIC回归系数	标准误	p值
截距项	-6.120	7.739	0.435
t日累计确诊增长率	3.083	0.562	<0.001
t日累计治愈增长率	-0.496	0.297	0.105
t日累计密切接触增长率	-1.614	0.608	0.013
F检验	p值<0.001		
调整后的 R^2	0.633		

t日累计确诊增长率正显著

- 累计确诊增长率每增加1个单位， $t+1$ 日新增确诊增长率平均增加3.083个单位。
- 累计确诊增长率越高，形势越严峻，因此下一日新增确诊增长率也可能走高。

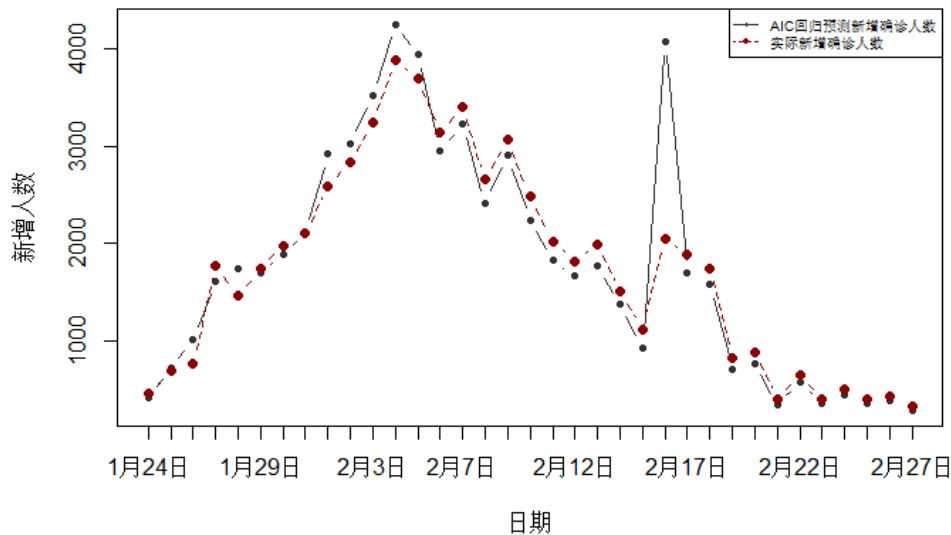
t日累计密切接触增长率负显著

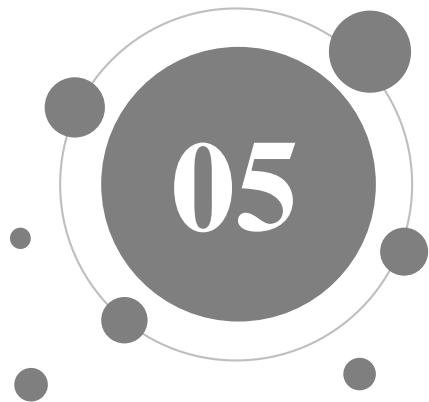
- t日密切接触增长率每增加1个单位， $t+1$ 日新增确诊增长率平均降低1.614个单位。
- 密切接触者人群范围的扩大是对疑似病例进行全面追踪的结果，因此结果反映了我国疫情防控措施的有效性。

模型预测

基于本案例的AIC模型，可以实现对未来一天的新增确诊增长率进行预测的目的

- 进行**样本内**预测，预测结果和实际结果较为吻合
- 对2月28日进行**样本外**预测，预测新增确诊人数为383，其95%的置信区间为334~434
- 2月28日真实的新增确诊人数为427，位于置信区间以内，说明AIC模型可以对未来一天的新增确诊人数进行较为精准的预测





PART FIVE

结论与展望

结论

- 8个增长率指标在因变量的高低水平组内均呈现明显差异
- 累计确诊增长率和累计密切接触增长率2个变量对 $t+1$ 日新增确诊增长率有显著影响
- 基于模型结果可以对 $t+1$ 日新增确诊人数进行较好的预测

局限性及展望

- 未考虑境外输入情况，随境外疫情的扩散，我国的境外输入病例也有所增加
- 未区分湖北省内外，湖北省疫情较其他省市而言更为严峻
- 未考虑无症状感染者的情况，无症状感染者可能对于 $t+1$ 日新增确诊增长率有重要的影响

狗熊会 | 精品案例



谢谢观赏！

扫描二维码，关注狗熊会，获取更多案例资源