



防护措施下一类高传染性多潜伏期传染病的高维动态模型

谈树萍 张纪峰

(中国科学院数学与系统科学研究院, 北京 100080)

摘要: 传染病给人们的生命带来了极大的威胁, 对于高传染性的疾病, 政府总会采取一些防护措施. 本文针对防护措施下的高传染性且具有潜伏期等特性的一类传染病, 结合传染病模型, 在一定假设条件下给出了这类疾病单日新收治的直接确诊病例及疑似病例的高维动态模型, 并使用最小二乘法进行了参数辨识. 最后以 SARS 为例, 利用网上公布的 SARS 数据给出了 5 月 22 日 - 5 月 31 日的预测结果并将预测结果和实际数据进行了比较, 说明了模型的有效性.

关键词: 传染病; 高维动态模型; 参数估计; 最小二乘法; 预测

中图分类号: O175.7; O212.1 **MR 分类号:** 92D25; 33E30 **文献标识码:** A

文章编号: 1001-9626(2005)04-0392-07

0 引言

虽然科技在不断发展, 但疾病总是威胁着人们的生命, 尤其是高传染性且具有潜伏期的疾病, 由于其具有潜伏期, 不仅使患者延误了最佳治疗时机, 而且使患者接触过的人在毫无防范的情况下被传染. 同时, 人们的流动性又使得这类高传染性疾病不断扩散到整个社会, 引起整体人群的恐慌, 对社会秩序、经济也造成一定影响. 当政府察觉到这类疾病的高传染性时, 会采取相应的措施切断传染链, 逐步控制疾病的传染. 虽然人们相信在政府的有效措施下, 疾病最终会被控制住, 但是到底什么时候会结束, 就很难给出明确的答案了, 人们依然生活在恐慌之中. 在这种情况下, 建立防护措施下这类高传染性多潜伏期传染病的高维动态模型, 预测今后疫情的走势就具有很重要的意义了. 一般的传染病模型没有给出延迟的情况, 即多潜伏期时的情况, 也没有考虑政府防护措施生效后的情况 [1,2], 本文针对这类高传染性且多潜伏期疾病的特性, 给出了政府防护措施下疾病的高维动态模型.

本文将传染病流行范围内的人群分成四类:

S 类: 新收治的疑似病例.

I 类: 新收治的直接确诊病例, 指直接被确诊为患者的, 不包括由疑似病例转为的临床诊断病例.

收稿日期: 2003-06-10

基金项目: 国家自然科学基金和科技部重大基础研究前期研究专项资助项目

作者简介: 谈树萍 (1978-), 女, 山东胶州人, 博士研究生.

R 类: 新增迁出者, 包括新收治的直接确诊病例、新收治的疑似病例, 新发现的与疑似或确诊病例接触过的、需要隔离的人群, 以及新增加的为所有疑似、确诊或其他被隔离的人员服务的医务人员.

E 类: 易感者, 指未被迁出、也未被收治、但与感病者接触后容易受到感染者.

1 建模及分析

结合高传染性疾病的实际情况, 相关的研究结果以及传染病模型, 本文提出以下假设:

H1 除感病特征外, 易感人群的个体间没有任何差异, 新收治的疑似病例与新收治的确诊病例的个体在人群中的混合是均匀的. 而且单位时间内每一个新收治的确诊病人所接触的易感人群占全部易感人群的比率, 称为接触率, 随政府采取的措施逐渐生效而开始衰减, 之后成为常数.

H2 易感人群感病的机会与他接触确诊病人的机会成正比. 这个比率, 称为感病率, 随政府采取的措施逐渐生效而开始衰减, 之后成为常数.

H3 易感人群转为疑似的机会与他接触确诊病人的机会成正比, 这个比率, 称为疑似率, 随政府采取的措施逐渐生效而开始衰减, 之后成为常数.

H4 从疑似病例转为确诊病例的比率为常数, 称为转移率.

H5 易感人群与新收治的确诊病人或新收治的疑似病人接触后被隔离的比率, 称为隔离率, 为常数.

H6 潜伏期从 a_1 天到 a_2 天不等.

H7 不考虑出生与死亡的过程.

由假设 H1, 第 n 天新收治的疑似病例和新收治的确诊病例的人数为 n 的函数, 分别记为 $S(n)$, $I(n)$. 第 n 天的迁出者记为 $R(n)$, 第 n 天的易感人群记为 $E(n)$, 记整体人群为 N . 接触率、感病率、疑似率分别记为 $K_1(n)$, $K_2(n)$, $K_3(n)$, 转移率、隔离率分别记为 α , β , 潜伏期记为 p , $p = a_1, a_1 + 1, \dots, a_2$, 其中 a_1, a_2 为正整数.

注 1 根据 $E(n)$ 的定义及假设 H1, 第 n 天的易感人群 $E(n)$ 指第 n 天还未被迁出, 也未被收治者, 与感病者接触后容易受到感染, 并且除感病特征外, $E(n)$ 个体间没有任何差异. 那么第 n 天的易感人群 $E(n)$ 应该是从整体人群中除去第 n 天之前所有的迁出者, 因为这些迁出者由于已经迁出而无法接触到感病者, 从而不属于易感人群; 并且除去第 n 天新收治的疑似病例和新收治的确诊病例的人数. 根据以上分析, 有

$$E(n) = N - I(n) - S(n) - \sum_{j=1}^{n-1} R(j).$$

同理第 $n-p$ 天的易感人群

$$E(n-p) = N - I(n-p) - S(n-p) - \sum_{j=1}^{n-1-p} R(j), \quad p = a_1, a_1 + 1, \dots, a_2.$$

分析假设 H1. 接触率 $K_1(n)$ 随政府采取的措施逐渐生效而开始衰减, 之后成为正的常数. 在政府采取的措施生效初期, 虽然接触率 $K_1(n)$ 已经开始衰退, 但由于这时 $I(n)$ 和 $S(n)$ 的值

还是很大, 所以从新收治的传染病确诊及疑似病例人数上看效果不很明显. 本文研究效果明显显现出的情形. 假设政府采取措施 m_1 天后, $K_1(n)$ 保持为常数 μ_1 . 同样的分析, 假设政府采取措施 m_2 天后, 感病率 $K_2(n)$, 疑似率 $K_3(n)$ 分别保持为常数 μ_2, μ_3 . 由于本文只研究防护措施生效后的模型, 不妨从第 m 天计时, $m = \max\{m_1, m_2\}$. 以下的 n 均代表防护措施生效后的第 n 天.

由于政府采取了隔离措施, 与当天出现的新收治的直接确诊病例接触过的易感人群以及与当天出现的新收治的疑似病例接触过的易感人群以 β 的隔离率被隔离, 从而当潜伏期过后, 在遗漏人群中可能有新收治的确诊病例以及新收治的疑似病例出现. 由于只有与症状期病人接触后才可能被感染, 从而我们认为潜伏期过后出现的所有的确诊病例以及疑似病例仅可能从这部分被遗漏的易感人群中出现. 换言之, 第 n 天新收治的直接确诊病例仅可能从 p 天前这部分被遗漏的易感人群中出现, $p = a_1, a_1 + 1, \dots, a_2$. 又注意到假设 H7, 由以上分析, 进而假设每天未被隔离的但已经感染病毒的人群以 ϕ_p 的概率在 p 天后成为新收治的直接确诊病例. 注意到假设 H1-H6, 对 $\forall n > a_2$, 对新收治的直接确诊病例有

$$I(n) = \sum_{p=a_1}^{a_2} \phi_p \{I(n-p)\mu_1\mu_2E(n-p)(1-\beta) + S(n-p)\alpha\mu_1\mu_2E(n-p)(1-\beta)\}. \quad (1)$$

同样的分析, 对新收治的疑似病例有

$$S(n) = \sum_{p=a_1}^{a_2} \phi_p \{I(n-p)\mu_1\mu_3E(n-p)(1-\beta) + S(n-p)\alpha\mu_1\mu_3E(n-p)(1-\beta)\}. \quad (2)$$

对迁出人群有

$$R(n) = I(n) + I(n)K_1(n)E(n)\beta + S(n) + S(n)K_1(n)E(n)\beta. \quad (3)$$

从 (1)-(3) 式看到, 单日新收治的确诊及疑似病例数与当日的易感人群, 确诊及疑似病例数有关, 假设在政府防护措施下累积迁出人群 $\sum R(n)$ 在整体人群 N 中所占比例不大, 可以不考虑普通人群 $E(n)$ 的变化, 认为 $E(n)$ 恒定, 为整体人群 N .

记

$$\lambda_p = \phi_p(1-\beta)N\mu_1\mu_2, \quad (4)$$

$$\sigma_p = \phi_p(1-\beta)N\mu_1\mu_3, \quad (5)$$

$$\varsigma = \beta N\mu_1. \quad (6)$$

根据参数 $\phi_p, \alpha, \beta, N, \mu_1, \mu_2, \mu_3$ 的含义, 参数 λ_p 表示整体人群中与一例直接确诊病人接触后未被隔离而且潜伏期 p 天过后成为直接确诊病例的人数, 参数 σ_p 表示整体人群中与一例直接确诊病人接触后未被隔离而且潜伏期 p 天过后成为疑似病例的人数, 参数 ς 表示整体人群中与一例直接确诊病人接触后被隔离的人数.

在新的参数下, 有

$$I(n) = \sum_{p=a_1}^{a_2} \{\lambda_p I(n-p) + \alpha\lambda_p S(n-p)\}, \quad (7)$$

$$S(n) = \sum_{p=a_1}^{a_2} \{ \sigma_p I(n-p) + \alpha \sigma_p S(n-p) \}, \tag{8}$$

$$R(n) = I(n) + \varsigma I(n) + S(n) + \varsigma S(n). \tag{9}$$

注 2 当 $a_1 = a_2 = c$, c 为某个正整数时, 即潜伏期为常数的情况.

2 参数辨识

本文应用最小二乘法辨识参数.

对 (7)-(9) 式, 应用最小二乘法辨识参数 λ_p 和 σ_p . 下面, τ 表示向量或矩阵的转置, 记

$$\lambda = (\lambda_{a_1}, \lambda_{a_1+1}, \dots, \lambda_{a_2})^\tau, \tag{10}$$

$$\sigma = (\sigma_{a_1}, \sigma_{a_1+1}, \dots, \sigma_{a_2})^\tau, \tag{11}$$

并把使

$$\sum_{i=a_2+1}^{n_0} \left\{ \sum_{p=a_1}^{a_2} [\lambda_p I(i-p) + \alpha \lambda_p S(i-p)] - I(i) \right\}^2$$

达到最小值的 λ 和 α 分别作为参数 λ 和 σ 的估计值, 并分别记作 $\hat{\lambda}$ 和 $\hat{\alpha}$, 即

$$(\hat{\lambda}^\tau, \hat{\alpha})^\tau = \arg \min \sum_{i=a_2+1}^{n_0} \left\{ \sum_{p=a_1}^{a_2} [\lambda_p I(i-p) + \alpha \lambda_p S(i-p)] - I(i) \right\}^2. \tag{12}$$

同理, 有 σ 和 ς 的估计值:

$$\hat{\sigma} = \arg \min \sum_{i=a_2+1}^{n_0} \left\{ \sum_{p=a_1}^{a_2} [\sigma_p I(i-p) + \hat{\alpha} \sigma_p S(i-p)] - S(i) \right\}^2, \tag{13}$$

$$\hat{\varsigma} = \arg \min \sum_{i=1}^{n_0} [I(i) + \varsigma I(i) + S(i) + \varsigma S(i) - R(i)]^2. \tag{14}$$

这里 n_0 为防护措施生效后的第 n_0 天.

求解 (13)-(14), 解得

$$\hat{\sigma} = (X_1^\tau X_1)^{-1} X_1^\tau Y_1, \tag{15}$$

$$\hat{\varsigma} = (X_2^\tau X_2)^{-1} X_2^\tau Y_2, \tag{16}$$

其中

$$X_1 = \begin{pmatrix} I(a_2 + 1 - a_1) + \hat{\alpha} S(a_2 + 1 - a_1) & \cdots & I(1) + \hat{\alpha} S(1) \\ \vdots & \vdots & \vdots \\ I(n_0 - a_1) + \hat{\alpha} S(n_0 - a_1) & \cdots & I(n_0 - a_2) + \hat{\alpha} S(n_0 - a_2) \end{pmatrix}, \tag{17}$$

$$Y_1 = \begin{pmatrix} S(a_2 + 1) \\ \vdots \\ S(n_0) \end{pmatrix}; \tag{18}$$

$$X_2 = \begin{pmatrix} I(1) + S(1) \\ \vdots \\ I(n_0) + S(n_0) \end{pmatrix}, \quad (19)$$

$$Y_2 = \begin{pmatrix} R(1) - I(1) - S(1) \\ \vdots \\ R(n_0) - I(n_0) - S(n_0) \end{pmatrix}. \quad (20)$$

当各个参数被估计出后, 就可以利用当日的数据预测潜伏期后的新收治的确诊病例及疑似病例.

3 实例结果及分析

3.1 模型下的预测结果

下面以 SARS 为例, 说明模型的有效性. 本文所给的单日新收治的确诊病例及疑似病例模型是政府采取的措施明显生效后的模型, 由于一般认为北京市政府自 4 月 20 号以后开始针对 SARS 采取大规模的防护措施, 结合中科院管理与控制 SARS 研究小组的结果 [3], 假定采取措施十天后 $K_1(n), K_2(n), K_3(n)$ 都达到最小值, 所以取 $m = 11$, 那么自 4 月 21 日开始第 11 天为 5 月 1 日. 根据网上公布的 5 月 1 日到 5 月 21 日之间北京市单日新收治的直接确诊病例及疑似病例的数据 (见附表), 对参数进行辨识. 由于本文将非典病区的医护人员归为 R 类, 所以需要对网上公布的单日新收治的 SARS 直接确诊病例和疑似病例数据进行处理. 假设医护人员中新增确诊病例全部是由疑似病例转为的临床确诊病例. 由于中国疾病预防控制中心公布非典病人的潜伏期一般为三到四天, 分别取 $a_1 = 3, a_2 = 4$ 讨论.

修正数据后利用网上数据辨识模型中的参数, 得:

$$\hat{\alpha} = 0.08; \quad (21)$$

$$\hat{\lambda}_3 = 0.24, \hat{\lambda}_4 = 0.23; \quad (22)$$

$$\hat{\sigma}_3 = 0.50, \hat{\sigma}_4 = 0.81. \quad (23)$$

利用辨识出的参数及动态模型可以预测出 5 月 22 日 - 5 月 31 日单日新收治的直接确诊及疑似病例. 结果如下:

表 4.1.1 北京市 5.22-5.31 日单日新收治的直接确诊病例及疑似病例预测

Table 4.1.1 Daily Numbers of SARS Patients and of SARS Suspect Patients in Beijing from May 22 to 31

日 期	5.22	5.23	5.24	5.25	5.26	5.27	5.28	5.29	5.30	5.31
新收治的直接确诊病例	4	2	2	2	2	1	1	1	1	1
新收治的疑似病例	12	6	5	5	3	3	3	2	2	2

如下的图 4.1.2 为 5 月 22 日 - 5 月 31 日预测值与实际值的比较, 其中实线表示实际值, 虚线表示预测值:

进一步的, 可以利用本文所给模型以及新获得的数据不断调整参数, 进行逐日滚动预测. 下面的图 4.1.3 为利用新数据逐步预测所得值与实际值的比较, 其中实线表示实际值, 虚线表

示预测值:

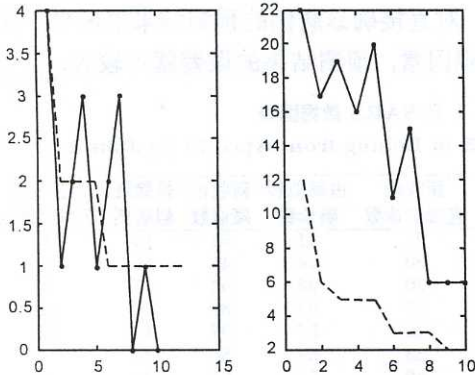


图 4.1.2 5 月 22 日 -5 月 31 日预测值与实际值的比较

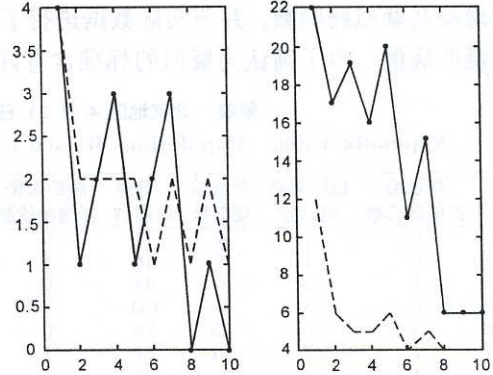


图 4.1.3 5 月 22 日 -5 月 31 日逐步预测值与实际值的比较

3.2 分析

根据参数 $\lambda_3, \lambda_4, \sigma_3, \sigma_4$ 的含义, 及辨识后所得的参数值知, 整体人群中与一例直接确诊病人接触后未被隔离的人中, 潜伏期过后成为直接确诊病例的人数较成为疑似病例的人数要少, 这在一定程度上反映出人们提高了对传染病的警惕性. 这一结论和每天都有大量的排除疑似病例是相符的. 由表 4.1.1. 的预测结果可知, 北京市 5 月 22 日 -5 月 31 日的单日新收治的直接疑似病例并不是单调下降的, 会不断出现波动, 但总体呈现下降趋势. 这一点是和事实吻合的. 从图 4.1.2 中可以明显的看出, 模型所给的单日新收治的直接确诊病例预测结果与实际数据的误差均值为 0, 而且绝对误差均值也不大, 经计算为 0.8 人, 这一结果非常理想. 但是单日新收治的疑似病例预测结果与实际数据的误差均值相对要大很多, 预测值较实际值要每天要少 9 人. 这也是和实际相符的. 由附表中的数据可以发现, 每天都有很多的疑似排除病例, 尤其是到了五月底, 每天的排除疑似病例远远超过了当天的新增疑似病例, 这必定会对疑似病例的预测带来影响. 由于单日新收治的直接确诊病例基本不存在这样的问题, 所以对单日新收治的直接确诊病例的预测精度相对就很高了. 而且, 由图 4.1.3 可见如果用新的数据不断调整参数, 进行逐日滚动预测, 预测结果会更加理想.

不过, 以上结果可能和实际有所偏差. 一是, 本文依一般情况假定 SARS 的潜伏期为三、四天, 这个数字在统计意义上是有效的, 但是当疫情衰退时, 由于病例数很小, 假定潜伏期为三至四天对感病个体就可能有一定出入, 从而使得预测数字和实际有偏差. 二是, 本文假定政府采取措施 10 天以后, 感染率, 接触率等衰减到常数. 但实际中, 人们可能因为单日新收治的直接确诊及疑似病例呈衰减态势而放松警惕, 使得感染率和接触率重新回升, 从而使疫情出现波动, 收尾减缓, 甚至出现反复. 三是, 每天有很多疑似排除病例, 一定程度上影响了疑似病例的预测.

4 结束语

本文在传染病模型基础上, 结合高传染性多潜伏期传染病的特殊性, 及政府的防护措施产生的重大作用, 给出了防护措施生效后单日新收治的直接确诊及疑似病例的动态模型, 并以

SARS 为例根据网上公布的数字辨识出参数值, 预测了 5 月 22 日-5 月 31 日的单日新收治的直接确诊及疑似病例数, 并与实际数据进行了比较, 对直接确诊病例的预测结果很理想. 但是对于疑似病例, 由于确认为疑似的标准含有许多主观因素, 预测结果的误差就比较大.

附表 北京地区 4 月 21 日-6 月 1 日 SARS 数据图表

Appendix Table Populations Related to SARS in Beijing from April 21 to June 1

日期	新收治 直接确诊数	由疑似转 确诊数	新收治 疑似数	排除疑 似病例	医护人员 新增确诊数	日期	新收治 直接确诊数	由疑似转 确诊数	新收治 疑似数	排除疑 似病例	医护人员 新增确诊数
6.1	0	1	4	11	0	5.11	21	21	51	16	4
5.31	0	1	6	18	1	5.10	30	24	45	49	4
5.30	1	5	6	44	0	5.9	20	28	54	87	3
5.29	0	3	6	141	0	5.8	37	57	80	51	12
5.28	3	0	15	79	0	5.7	40	57	97	49	12
5.27	2	7	11	11	1	5.6	33	37	80	30	14
5.26	1	4	20	52	1	5.5	40	58	65	34	9
5.25	3	10	16	35	1	5.4	43	26	90	20	8
5.24	2	24	19	40	1	5.3	66	48	104	31	20
5.23	1	14	17	29	1	5.2	48	48	145	44	15
5.22	4	11	22	27	1	5.1	82	40	96		20
5.21	1	7	34	31	0	4.30	70	31	116	18	
5.20	2	10	17	31	1	4.29	103	49	149	50	
5.19	3	4	27	24	2	4.28	59	37	149	20	
5.18	9	8	32	24	1	4.27	76	50	162	20	
5.17	6	13	41	43	2	4.26	113		173	7	
5.16	6	22	43	73	2	4.25	103		161	17	
5.15	15	12	41	20	2	4.24	89		112	17	
5.14	21	18	51	63	3	4.23	105				
5.13	14	34	35	41	4	4.22	106				
5.12	9	39	48	42	5	4.21					

参 考 文 献

- [1] [专著] 刘来福, 曾文艺. 数学模型与数学建模. 第二版 [M]. 北京: 北京师范大学出版社, 1997, 150-1997.
 [2] [专著] (美) W.F. 卢卡斯主编; 翟晓燕, 黄振高, 许若宁译, 生命科学模型 [M]. 长沙: 国防科技大学出版社, 1996, 252-262.
 [3] 龚其同, 佟仁城, 吕本富等. 北京市 SARS 的防疫措施和疫情的发展趋势分析 [J]. 管理评论, 2003, 15(4):28-30.

A High-Dimensional Propagation Model for a Class of Highly Infectious Latent Diseases under Safety Measures

TAN Shu-ping ZHANG Ji-feng

(Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100080 China)

Abstract: Considering the great threat from infections against people's lives, safety measures are always taken to avoid spreading of highly infectious diseases. In this paper, with respect to a class of highly infectious latent diseases under safety measures, a high-dimensional dynamical model is presented for the daily populations of patients and of suspect patients based on the concepts of previous epidemic model and some mild assumptions. Least square algorithm is used to estimate the model parameters. Finally, as an example, the daily number of SARS patients and that of SARS suspect patients in Beijing from May 22 to 31 is predicted based on the data opened on network. Comparison of the prediction data with the real data shows the efficiency of our model.

Key words: Infection ; High-dimensional dynamical model ; Parameter estimation ; Least square algorithm ; Prediction