

吉林市 2010 年丙型病毒性肝炎发病率的灰色系统 GM(1,1) 模型预测

程志勇, 赵贺春 (吉林市疾病预防控制中心, 吉林 吉林 132001)

【摘要】 目的 利用灰色系统 GM(1,1) 模型对吉林市 2010 年丙型病毒性肝炎发病率进行预测, 为丙型肝炎的防治提供科学依据。方法 利用 1999—2009 年吉林市丙型肝炎发病率统计资料建立灰色系统 GM(1,1) 模型, 并对模型进行评价。结果 丙型肝炎发病率预测模型为 $Y_t = 9.967\ 02 e^{0.267\ 73(t-1)} - 8.351\ 32$ 模型检验精度好, 外推 2010 年发病率为 44.488 4/10 万, 检验效果满意。结论 利用灰色系统 GM(1,1) 模型能够对吉林市 2010 年丙型肝炎做出合理预测。

【关键词】 丙型肝炎; 灰色系统; GM(1,1) 模型; 预测
〔中图分类号〕 R195 R181 〔文献标识码〕 A 〔文章编号〕 1671-4199(2011)02-0158-02

传染病预测是根据传染病的发生、发展规律及相关因素, 用分析判断和数学模型等方法对可能发生的传染病的发生、发展和流行趋势作出预测, 是制定防控传染病的长远或近期应对策略的前提。本研究利用灰色系统 GM(1,1) 模型对 2010 年吉林市丙型肝炎 (以下简称丙肝) 进行预测, 现报告如下。

1 资料与方法

1.1 资料来源 1999—2009 年吉林市法定传染病疫情监测丙肝发病率数据 (按发病日期统计) 统计资料。

1.2 方法和原理

1.2.1 灰色系统 GM(1,1) 模型的基本思想 将无规律的原始数据累加生成后, 使其变为有规律的生成数列, 然后建立相应的微分方程模型, 从而预测事物未来发展趋势和状态。

1.2.2 GM(1,1) 模型的建模过程

1.2.2.1 累加生成 设原始数列 $X_t(t=1, 2, \dots, n)$, 对其进行一次累加生成数列 Y_t 即:

$$Y_t = \sum_{i=1}^n X_i (t=1, 2, \dots, n) \tag{1}$$

1.2.2.2 均值生成 对累加生成数列 Y_t 作均值生成, 即

$$Z_t = \frac{1}{2} (Y_t + Y_{t-1}), (t=2, 3, \dots, n) \tag{2}$$

1.2.2.3 建立 GM(1,1) 模型 Y_t 的估计值 \hat{Y}_t 其一阶线性微分方程为:

$$\frac{dY_t}{dt} + \alpha Y_t = \mu \tag{3}$$

1.2.2.4 解微分方程 式 (3) 中 α 和 μ 是待定参数, 其中 α 为内生灰数, μ 为发展灰数, 根据最小二乘法估计参数, 建立数据矩阵 B

$$B = \begin{bmatrix} -Z_2 & 1 \\ -Z_3 & 1 \\ \vdots & \vdots \\ -Z_n & 1 \end{bmatrix}, \text{ 则有 } a = \begin{pmatrix} \alpha \\ \mu \end{pmatrix} = \frac{B^T Y}{(B^T B)} \tag{4}$$

通过矩阵运算可求得矩阵 a 中待定参数 α 、 μ 值。

解微分议程式 (3) 得:

$$Y_t = (X_1 - \frac{\alpha}{\mu}) e^{-\alpha(t-1)} + \frac{\mu}{\alpha} \tag{5}$$

1.2.2.5 估计值的计算

$$\hat{X}_t = \hat{Y}_t - Y_{t-1} (\hat{X}_t \text{ 为 } X_t \text{ 估计值}) \tag{6}$$

表 1 拟合精度等级表

精度等级	P 值	C 值
好	> 0.95	< 0.35
合格	> 0.80	< 0.45
勉强合格	> 0.70	< 0.50
不合格	≤ 0.70	≥ 0.65

1.2.3 模型的检验 设残差 $\delta_t = X_t - \hat{X}_t$, $\bar{\delta}$ 为原始数据序列的标准差, $\bar{\delta}$ 为残差的标准差。

$$\bar{\delta} = \sqrt{\frac{\sum_{t=1}^n (X_t - \bar{X})^2}{N}} \text{ 其中 } \bar{X} \text{ 为原数列均值} \tag{7}$$

$$\bar{\delta} = \sqrt{\frac{\sum_{t=1}^n (\delta_t - \bar{\delta})^2}{N}} \text{ 其中 } \bar{\delta} \text{ 为残差的均值} \tag{8}$$

作者简介: 程志勇 (1972—), 男, 吉林吉林人, 学士, 主管医师, 研究方向: 传染病预警预测。

计算后验差比值 $C=\xi/\hat{\xi}$ (9)
小误差概率为 $P\{|\hat{\alpha}-\bar{\alpha}|\leq 0.6745\hat{\sigma}\}$ (10)
根据 C值及 P值的大小,综合评价模型的拟合效果,常用的拟合精度等级见表 1。

1.3 统计学处理 应用统计软件 Excel2003 进行分析。

2 结果

2.1 吉林市 1999—2009年丙肝发病率 数列为 $X=(1.675\ 7, 2.862\ 7, 4.703\ 3, 5.352\ 3, 6.461\ 4, 10.670\ 3, 12.193\ 0, 14.966\ 3, 20.171\ 6, 27.136\ 8, 35.938\ 3)$, 设其时序分别为 1, 2, 3... 11, 按公式 (1)、(2)进行累加生成和均值生成,得到数列 Y 和 Z , 见表 2

表 2 1999—2009年吉林市丙肝发病率 (1/10万) 的 GM(1, 1)模型建模数据表

年度	时序	发病率 X_t	Y_t	Z_t
1999	1	1.675 7	1.675 7	
2000	2	2.862 7	4.538 4	3.107 1
2001	3	4.703 3	9.241 7	6.890 1
2002	4	5.352 3	14.594 0	11.917 9
2003	5	6.461 4	21.055 4	17.824 7
2004	6	10.670 3	31.725 7	26.390 6
2005	7	12.193 0	43.918 7	37.822 2
2006	8	14.966 3	58.885 0	51.401 9
2007	9	20.171 6	79.056 6	68.970 8
2008	10	27.136 8	106.193 4	92.625 0
2009	11	35.938 3	142.131 7	124.162 6

根据公式 (4) 计算 $\alpha = -0.267\ 73$ $\mu = 2.235\ 90$ 根据公式 (5)则预测模型公式为

$$Y_t = (1.675\ 7 - \frac{2.235\ 90}{-0.267\ 73})e^{267\ 73(t-1)} + (\frac{2.235\ 90}{-0.267\ 73})$$

即为: $Y_t=9.967\ 02e^{267\ 73(t-1)}-8.351\ 32$ (11)

2.2 模型检验 将 $t=1, 2\cdots 11$ 代入公式 (11)所得各时序 Y_t 值,然后按公式 (6)还原得各时序模型拟合值 X_t 并计算残差如表 3

根据公式 (7) - (10)计算 $\hat{\xi}=10.901\ 4$ $\hat{\xi}=0.787\ 5$ 后验差值比 $C=0.722$ 小概率误差 $P=1$, 对照表 1,精度为好,可进行外推预测。

2.3 模型外推预测 将 $t=12$ 代入模型 $Y_t=9.967\ 02e^{267\ 73(t-1)}-8.351\ 32$ 对 2010年丙肝预测, $Y_{12}=181.071\ 5$ 由 $Y_{12}=Y_{12}-Y_{11}=181.071\ 5-142.131\ 7=44.488\ 4$ 2010年吉林市丙肝预测值为 44.488 4/10万。

表 3 1999—2009年吉林市丙肝发病率 (1/10万) 的 GM(1, 1)模型残差表

年份	时序	发病率 (X_t)	预测值 \hat{X}_t	残差 $\hat{\alpha}$
1999	1	1.675 7	1.615 7	0
2000	2	2.862 7	3.059 4	-0.196 7
2001	3	4.703 3	3.998 6	0.704 7
2002	4	5.352 3	5.225 9	0.126 4
2003	5	6.461 4	6.830 0	-0.368 6
2004	6	10.670 3	8.926 5	1.743 8
2005	7	12.193 0	11.666 6	0.526 4
2006	8	14.966 3	15.247 7	-0.281 4
2007	9	20.171 6	19.928 0	0.243 6
2008	10	27.136 8	26.045 0	1.091 8
2009	11	35.938 3	34.039 7	1.898 6

3 讨论

灰色系统理论诞生于 1982年由我国学者邓聚龙教授创立,是一种研究少数据、贫信息不确定性问题的新方法。灰色系统理论以“部分信息已知,部分信息未知”的“小样本”、“贫信息”不确定性系统为研究对象,主要通过对“部分”已知信息的生成、开发,提取有价值的信息,实现对系统运行行为、演化规律的正确描述和有效监控。

传染病的发病率受诸多因素的影响,人们对其发病因素的了解常常是不完全的,其中有的已为人们所认识,有的还需要进一步研究探索。正是由于传染病发病或明或暗的灰色性,构成了一个灰色系统^[1]。灰色预测是基于微分方程的预测, GM(1, 1)模型是最基本的模型。它以独特的数学方法把难以描述的理论作为灰色理论来处理,弱化随机因素的干扰,从杂乱无章的现象中揭示事物的发展规律,对样本容量和概率分布没有严格要求,模型简单,预测效果好,适合于对流行因素较稳定的疾病进行短期预测^[2];应用 Exce软件即可编写模型预测模板^[3], 便于在基层疾病预防控制机构推广。

笔者利用 1999—2009年 11年吉林市丙肝发病数据,建立灰色系统 GM(1, 1)模型,通过模型检验,效果评价为好,并通过模型外推,对 2010年丙肝发病趋势进行预测,为吉林市丙肝防治规划调整及防治效果评价提供数据参考。

4 参考文献

[1] 易静,杜昌廷,王润华,等.应用灰色预测模型 GM(1, 1)对结核病发病率进行预测[J].重庆医科大学学报, 2007, 32(3): 275—278.
[2] 邢慧娟,杨维中,王汉章,等.传染病预测[J].预防医学情报杂志, 2004 20(6): 639—641.
[3] 陈青山,王声滂,迟桂波,等.应用 EXCEL完成性病 GM模型的预测和评价[J].疾病控制杂志, 2003 7(5): 451—453.

[收稿日期 2010-12-09]

(编校 张弘/曲莉)