维普资讯 http://www.cqvip.com

论 著 文章编号:0253-3626(2008)04-0472-03

重庆市育龄妇女不孕率的趋势面分析

路 超[1],田考聪[1],陈和平[2]

(1. 重庆医科大学公共卫生学院统计教研室,重庆 400016;2. 重庆市人口和计划生育科学技术研究院,重庆 400020)

【摘 要】目的:研究重庆市育龄妇女不孕率的地理分布规律。方法:应用趋势面模型探讨重庆市各区县育龄妇女不孕率的空间分布趋势。结果:在东部地区不孕率由北向南逐渐增高,在西部地区则变为由南向北逐渐增高,东北部最低,东南和西北部最高。结论:重庆市育龄妇女不孕率呈明显的地域分布规律,为进一步探讨其地理危险因素提供了理论依据。

【关键词】不孕率;地理分布;趋势面分析

【中国图书分类法分类号】R195.4 【文献标识码】A 【收稿日期】2007-10-15

Trend-surface analysis of the geographical distribution of the rate of infertility in Chongqing city

LU Chao, et al

(*Department of Health Statistics, College of Public Health, Chongqing Medical University*)

【Abstract】 *Objective*: To study the distribution of the rate of infertility in Chongqing city. *Methods*: The geographical distribution of infertilities of different regions in Chongqing city was investigated based on the methods of the trend-surface analysis. *Result*: In the east part of Chongqing city the rate of infertility is lower in north, in the west part the rate is lower in south, and the rate is lowest in the northeast. *Conclusion*: The infertility has geographical distribution trends in Chongqing city. Research should be performed on the difference of the environment and living habit.

【Key words】 Infertility; Geographical distribution; Trend-surface analysis

许多疾病的发生具有一定的地区分布特征,趋势面分析是以多元回归分析理论为基础的一种统计方法。它从整体出发,分析疾病的地理分布系统和局部分析。趋势分析图则是根据最小二乘法原理,剔除局部和随机变异的影响后求得的趋势面,较通常采用的疾病地区分布图更能准确地反映疾病地区分布总的变化趋势[1]。

趋势面模型作为一种常用的疾病空间分析工具,具有以下优点[2]:(1)无须假设人口分布均匀,只需以发病率、患病率、死亡率等指标做因变量建立模型,即可消除人口变异带来的干扰;(2)模型连续,符合大多数缓慢起伏变化的疾病空间分布模式;(3)"趋势性"可以涵盖"聚集性"和"稀疏性"。因为趋势面上的"高"、"低"处对应于"聚集"、"稀疏";而"趋势"还包括系统趋势和局部异常等地理流行病学信息;(4)趋势面分析可以预测各地的发病率或死亡率。

作者介绍:路 超(1981-),男,硕士,
研究方向:GIS 与统计的结合应用。
通讯作者:田考聪,男,教授,E-mail:tkc5155@126.com。

1 材料与方法

1.1 资料来源

由重庆市计生医院提供的 2005 年各区县育龄妇女不孕率及按重庆市地图确定各地观察点的地区位置。见表 1。

1.2 研究方法

根据各观察点的孕产妇不孕率及其相应地域位置,建立二元多项式回归进行趋势面分析。

1.2.1 趋势面方程的建立 设在观察点 $Z_i(x_i,y_i)(i=1,2,\cdots,n)$ 处孕产妇不孕率 $\hat{Z}_i(x_i,y_i)$,则趋势面方程形式为:

$$Z_i(x_i,y_i)=\hat{Z}_i(x_i,y_i)+\varepsilon_i$$

式中:ε_i 即为剩余值(残差值)。显然,当 (x_i,y_i) 在空间上变动时,上式就刻画了地理要素的实际分布曲面、趋势面和剩余面之间的互动关系。

多项式趋势面的形式 :

一阶趋势函数 $Z_1=a_0+a_1x+a_2y$

维普资讯 http://www.cqvip.com

表 1 重庆市不孕率观察点位置及不孕率

区县	经度(X)	纬度(Y)	不孕率(1/万)
铜梁	29.51	106.04	57.00
巴南	29.23	106.32	42.00
武隆	29.19	107.445	67.00
长寿	29.49	107.05	55.00
石柱	30	108.07	56.00
秀山	28.27	108.59	136.00
云阳	29.58	108.54	79.00
璧山	29.35	106.13	45.00
丰都	29.53	107.43	84.00
垫江	30.2	107.21	68.00
荣昌	29.23	105.35	49.00
江津	29.17	106.15	41.00
忠县	30.175	108.02	52.00

二阶趋势函数 $Z_2=a_0+a_1x+a_2y+a_3x^2+a_4xy+a_5y^2$

三阶趋势函数 $Z_3=a_0+a_1x+a_2y+a_3x_2+a_4xy+a_5y^2+a_6x^3+a_7x^2y+a_8xy^2+a_9y^3$

K 阶趋势函数 $Z_3=a_0+a_1x+a_2y+a_3x_2+a_4xy+a_5y^2+a_6x^3+a_7x^2y+a_8xy^2+a_9y^3+\cdots+a_py^k$

$p=1/2(k+1)(k+2)-1$

根据最小二乘法原理解方程组,得到趋势函数回归系数估计值:$a_1,a_2,a_3\cdots a_p$。将回归系数估计值代入方程即得相应阶次的趋势面方程。

1.2.2 趋势面模型的适度检验 趋势面分析拟合程度与回归模型的效果直接相关,因此,对趋势面分析进行适度性检验是一个关系到趋势面能否在实际研究中加以应用的关键问题,也是趋势面分析中不可缺少的重要环节。

趋势面拟合适度的 R^2 检验:趋势面与实际面的拟合度系数 R^2 是测定回归模型拟合优度的重要指标。一般用变量 z 的总离差平方和总回归平方和所占的比重表示回归模型的拟合优度。

$$R^2=\frac{SS_{回}}{SS_{总}}=1-\frac{SS_{剩}}{SS_{总}}$$

R^2 越大,趋势面的拟合度就越高。

1.2.3 选择趋势面方程 趋势面方程并不是阶次越高越好。阶次高的多项式虽然从逼近观察点数据方面效果较好,但其所表示的曲面过于复杂,有时反而偏离实际趋势,在内插或外推上效果不好。低次趋势面较适用于区域性变化的研究,高次趋势面则更能反映较小范围的变异。一般来说,应根据趋势面方程显著性检验结果、拟合优度和标准误的大小等,结合实际情况,综合考虑,选择适当阶次的趋势函数。

2 结果及绘制趋势图

在趋势值固定(等高)的条件下,设定 X_i,求出 Y_i,按照 X_i 和 Y_i 绘制趋势面分析图。本项研究数据是应用 SAS 8.2 软件包进行整理和分析的,并绘制了等高线图。

2.1 SAS 程序实例及结果

对表 1 数据资料作趋势面分析并作趋势图,SAS 程序如下:

```
/* 建立数据集 */
data trend;
input x y w;
label x='经度' y='纬度' w='不孕率(1/万)';
cards;
29.51    106.04    57
29.23    106.32    42
29.19    107.445   67
29.49    107.05    55
30.00    108.07    56
28.27    108.59    136
 ...      ...       ...
;
/* 建立趋势方程 */
proc glm data=trend;
model w = x y x*x x*y y*y /p;
run;
```

各阶趋势面拟合结果见表 2。

表 2 重庆市育龄妇女不孕率趋势面拟合结果

趋势面	F 值	P 值	$MS_{剩余}$	R^2
1 阶	14.23	0.0012	202.035	0.740012
2 阶	18.4	0.0007	78.480	0.929305
3 阶	18.51	0.0007	78.054	0.929690

由于 2 阶 R^2 与 3 阶 R^2 相差只有 0.000 385,但计算量却减少很多,故得出趋势方程最佳表达式:

w=−80 538.409 69+5 295.277 50×x+34.803 53×y−9.563 86×x×x−44.040 98×x×y+5.939 72×y×y

```
/* 图形线形选项设置 */
title font = '宋体' '重庆市 13 区县不孕率分布等高图';
axis1 label=none value=('28.00' '29.00' '30.00' '31.00')
      color=red width=3;
```

维普资讯 http://www.cqvip.com


```
axis2 label=none value=('105.00' '106.00'
'107.00' '108.00' '109.00')
        color=red width=3;
legend frame;
/* 用 GCONTOUR 过程绘制等高线等高图,结果见图 1 图 2*/
proc gcontour data=trend;
plot x*y=w/autolabel=(check=none)
            levels =40 to 400 by 40
              haxis=axis1
              vaxis=axis2
              legend=legend1;;
run;
proc gcontour data=trend;
 plot y*x=w / pattern
              coutline=gray
              ctext=green;
run;
```

2.2 图形输出结果

见图 1、2。

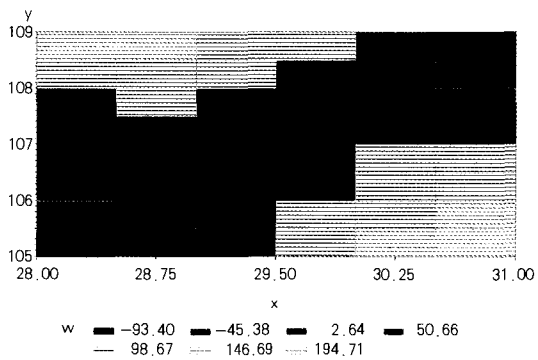


图 1 重庆市 13 区县不孕率分布等高图

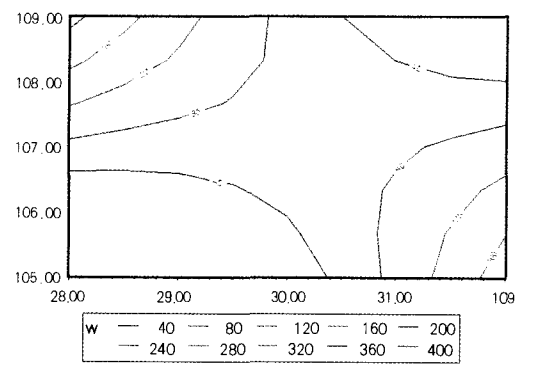


图 2 重庆市 13 区县不孕率分布等高线图

3 讨 论

疾病的地理分布特征往往不仅包括区域(或系统性)因素引起的变化,而且也包含着局部性因素和随机性因素所致的变异。趋势面分析法将每一观察值分解为趋势值和剩余值两部分。趋势值所组成的趋势面表示研究区域的系统性变异,这种变异一般被认为是由于环境总的变化或人群分布的系统性变化所致。剩余包括随机误差和局部因素所引起的变化。如果趋势面能完全代表观察值,则剩余是由随机误差所致。在实际工作中,拟合往往是不完全的,所以剩余既包括随机误差,也反映了局部范围的变异。趋势面分析不但能从参差不齐的数据中别除局部和随机变异的影响,显示大范围区域的系统性变化,研究总的变化规律,而且也能反映出疾病局部地区变异,为病因探讨提供更为重要的线索。

本文图 1 和图 2 所提供的信息和趋势基本一致,只是表现形式不同,图 2 用不同的线型对应 w 的不同水平值,图形清晰明了;而图 1 以不同的图案填充 w 的每个水平值区域,它能在局部提供更详细的信息,如图 1 中左下角的一小块水平值区域在图 2 中是无法体现的。两者各有优缺点,在应用时要根据具体情况对其进行选择。

本文所得到的趋势方程的拟合优度(即复相关系数的平方)R^2 为 0.898 783($P<0.001$),表明所取的趋势面方程能反映实际变异的 89.878 3%,可较好地描述重庆市不孕症发生率的地理分布规律性。研究结果显示,不孕症有明显的地区分布特征(见图 1~2)。提示地理环境,医疗卫生及经济等因素在不孕症发生中可能起着重要的作用。

参 考 文 献

[1] 温焕新,洪 楠.用趋势面分析法研究我国足月低体重儿的地理分布[J]. 中山医科大学学报,1991;12(4):306.

[2] 湖南医学院肿瘤研究室. 趋势面分析及其在肿瘤死亡回顾性调查中的应用[J]. 湖南医学院学报,1980;5(1):18-27.

(责任编辑:谭乐勤)