

应用对分法和样条法判定 昆虫种群空间分布型

周 国 法

(北京大学分校数学系)

李 天 生

(中国林业科学研究院林业研究所)

摘要 本文提出了判定昆虫种群空间分布型的样条法及对分法,并以马尾松毛虫为例进行了计算和分析;用模拟及实例说明了样条法和对分法判断分布型的依据。该方法能判断昆虫种群的空间分布型和聚集块的位置、大小及聚集程度;与已有的分布型指数相比,更能准确反映昆虫种群在空间的分布特征。

关键词 对分法;样条法;昆虫种群空间分布型

关于昆虫种群空间分布型的研究已有不少方法^[1,2],尤其是分布型指数法发展很快^[3]。分布型指数虽然可以判定昆虫的卵、幼虫(若虫)等的分布是聚集分布、随机分布还是均匀分布,但大多数分布型指数不能判定昆虫种群聚集块的范围。

本文提出使用对分法及样条法^[4],研究昆虫种群空间分布时,若为聚集分布,还可以进一步判定聚集范围的大小,由此,为研究昆虫种群空间分布型提供了更有效的手段。本文以马尾松毛虫种群调查资料为例说明对分法和样条法的计算方法。

1 对分法原理及计算实例

在作森林虫情调查时,将所研究的林区划分为正方形网格状^[5],网格大小依具体情况而定。记录各网格内的虫数,并写在相应位置的网格内,即每格一个原始数据,

网格划分时格数要求 2^h 个。将全部网格作对分,对分要求相邻的两行(列)合并(见图1),对合并后的网格再作列(行)合并,以此类推,直到最后合并为一个数据为止。

记 T_1 是全体数据平方和,第一次对分后数据平方和记为 T_2 ,第二次对分后数据平方和记为 T_4, \dots ,一般简单记为 $T_r, r=1, 2, 4, \dots, N=2^h$;又记 $G_r=2T_r-T_{2r}, r=1, 2, \dots, N/2$,则 $M_r=G_r/N$ 为总体样本方差的估计值。

在 N 一定时简单分析一下 G_r 的特点。实际上 G_1 (在水平对分中,下同)反映了相邻两行

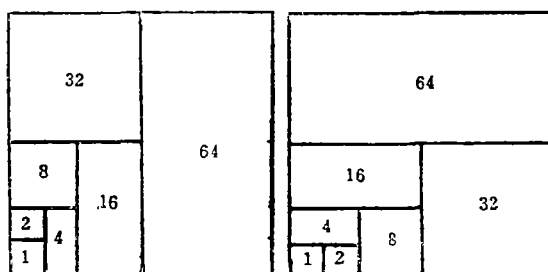


图1 对分法图示(左：竖直对分；右：水平对分)

间差别的大小， G_2 反映了相邻的四格间数据差别的大小， G_4 反映了相邻四行间数据差别的大小，其余类似。由此可知，若 G_r 有一个明显的峰值，则说明相邻的 r 格间虫口差别较大， G_r 越大，虫口差别越大，即存在聚集块。如果 G_r 没有明显的峰值，则不存在聚集块。因此可以通过 G_r 的值判断昆虫的空间分布型。具体方法如下：

(1) 若 G_r 有明显的峰值，则判定该分布为聚集分布，聚集范围正好是峰值处 r 对应的网格范围。

(2) 若 G_r 无明显峰值，则计算 G_r 与 $k(r = 2^{k-1})$ 间的相关系数 R 。

若 $R > 0$ ，判定该分布是随机的，若 $R < 0$ ，判定该分布是均匀的。

关于 G_r 是否有明显的峰值，可用 F -检验判断， $F_r = G_r/G_1$ ，查 $F(N/2r, N/2)$ 临界值 $F_{0.05}$ ，若 $F_r > F_{0.05}$ ，则 G_r 是显著的峰值(与 G_1 比较)。若同时存在多个 G_r 显著大于 G_1 ，若 G_{r_1} 是其中 r 值最小的，则若 $G_{r_2}(r_2 > r_1)$ 显著大于 G_{r_1} ，即认为聚集范围由 r_2 确定，否则仍由 r_1 确定。

下面以浙江省安吉县龙山林场，1983年马尾松毛虫越冬代幼虫，及1984年第一代幼虫调查资料为例，说明对分法计算方法及结果分析。将林区划分为 $1.7\text{ m} \times 1.7\text{ m}$ 网格，逐格调查虫口，部分调查数据如表1形式。

表1 第一代马尾松毛虫幼虫调查部分数据
(网格大小： $1.7\text{ m} \times 1.7\text{ m}$)

0	0	0	0	0	0	0	0	0	0	0	0	0	0	24	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	25	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	4	12	0	0
0	0	5	0	0	0	0	1	0	1	0	5	0	12	0	0	0
0	0	0	0	2	0	0	2	0	0	0	0	1	6	0	2	0
0	0	4	0	0	0	0	0	0	0	0	4	0	33	0	13	0
0	0	1	0	1	0	1	0	0	0	0	1	0	19	0	13	0
0	0	0	0	0	0	0	0	0	0	0	0	0	17	0	12	0
0	0	0	0	0	3	5	0	0	9	3	0	1	30	0	13	0
0	0	0	4	3	0	0	0	0	9	0	0	0	42	0	24	0
5	0	0	12	0	0	0	8	0	0	1	0	0	31	0	13	0
7	0	0	0	0	0	0	3	1	2	0	0	8	38	0	0	0
0	0	8	12	5	1	0	11	2	5	0	0	13	18	35	7	0
12	0	13	2	0	0	0	13	0	0	0	0	14	36	20	16	0
5	8	2	6	0	7	8	0	0	0	5	19	9	51	0	1	0
8	14	0	13	18	3	12	9	0	0	0	0	42	0	23	0	0

用上述方法对原始数据处理的结果见表2和表3。结果表明，第一代幼虫呈显著的聚集分布， $r=16$ 时 G_r 有一个明显的峰值($G_{16}/G_1 = 6.15 > 2.0$)，说明存在直径16格($16\text{ m} \times 1.7\text{ m}$)的聚集块。越冬代幼虫的聚集范围只有2格($G_2/G_1 > 1.38$)，从 M_r 容易看出，从 $r=4$ 到 $r=256$ 间 G_r 的值呈平稳上升趋势(相关系数 $R > 0$)，故从较大范围看，越冬代幼虫呈随机分布趋势。

2 样条法原理及计算实例

如作虫情调查时，对一个林区不作全部调查，只是沿着某些直线作网格调查，这时仍采用对分的方法处理数据称为样条法^[4,6]。它的调查类似平行线抽样法，样条的选取是

表2 第一代处理结果

r	T _r	G _r	M _r
1	21 813	16 393	64.0
2	27 233	8 686	33.9
4	45 780	15 271	59.7
8	76 289	5 823	22.8
16	146755	100 853	393.9*
32	192 657	22 515	88.0
64	362 799	114 728	448.1
128	610 873	290 521	1 134.9
256	931 225		

* 示 $\alpha=0.95$ 显著。

** 示与 $r=2$ 比较仍显著。

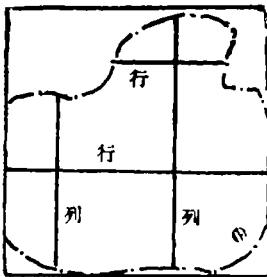


图2 样条法取样方法

$r=1, 2, 4, \dots$ 。

若 $N=2^h+m$, $m<N$, 先将原始数据作如下分组: $\{x_1, \dots, x_n\}, \{x_2, \dots, x_{n+1}\}, \dots, \{x_{m+1}, \dots, x_{m+n}\}$, $n=2^h$, 共 $m+1$ 组, 对每一组求 $G_r^{(l)}$, $l=1, 2, \dots, m+1$, 取平均

$$\bar{G}_r = \sum_{l=1}^{m+1} G_r^{(l)} / (m+1)$$

对 \bar{G}_r 采用对 G_r 的处理方法即可。

以越冬代和第一代松毛虫幼虫调查数据为例, 用样条法进行计算。共取三行、三列作调查, 数据较多的是通过林地中心地带的样条。

表4 样条法处理的越冬代幼虫结果

r	列			行		
	T _r	G _r	M _r	T _r	G _r	M _r
1	56	49	1.53	60	30	0.23
2	63	45	1.41	90	38	0.30
4	81	25	0.78	142	45	0.35
8	137	29	0.91	239	54	0.42
16	245	9	0.28	424	188	1.47*
32	441			660	580	4.53**
64				740	324	2.53
128				1 156		

* 超过95%的显著水平。

表3 越冬代处理结果

r	T _r	G _r	M _r
1	1 294	284	0.28
2	2 304	2 121	2.07*
4	2 487	760	0.74
8	4 214	3 034	2.96
16	5 394	4 986	4.87**
32	5 302	3 410	3.33
64	8 194	6 220	6.07
128	10 168	6 500	6.35
256	13 836	6 472	6.32
512	21 200	794	0.77
1 024	41 616		

随机的, 或根据某些事先已知信息采用的机械抽样, 如图2所示。

样条法的算法与对分法相同, 可看作一维时的对分法。设某样条调查的结果按地理位置顺序排列为 x_1, x_2, \dots, x_n , 计算为:

若 $N=2^h$, 则 $T_1 = \sum x_i^2$, $T_2 = \sum (x_{2i-1} + x_{2i})^2$, $T_4 = \sum (x_{4i-3} + x_{4i-2} + x_{4i-1} + x_{4i})^2$, \dots , 最后 $T_n = (\sum x_i)^2$, 同样 $G_r = 2T_r - T_{2r}$,

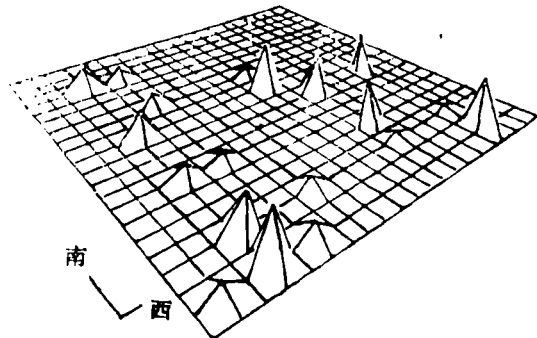


图3 越冬代幼虫的分布示意

表 4 结果说明,越冬代南北方向(列样条)呈均匀分布,因为 G_r-k 相关系数显著为负值。东西方向(行样条), $r=32$ 时 G_r 有一显著峰值,说明有一长 32 格的聚集块。综合上述结果知,在南北方向存在密度均匀的带状聚集块,见图 3。

由表 5 知第一代幼虫在列 1 及行 1 都存在直径 2 格的聚集块,而在林地中部(列 2、行 2)则存在长直径 8 格的聚块。

综上所述可得结论:越冬代幼虫在林中呈均匀分布,个别地方沿直线呈条状聚集。第一代幼虫呈带状聚集分布,该带状聚集在越冬代幼虫分布的直线周围(比较图 3 与图 4)。可见第一代幼虫的聚集块是由越冬代扩散形成的。

表 5 样条法处理的第一代幼虫结果

a. 一般处理

r	行 1			列 1			列 2		
	T_r	G_r	M_r	T_r	G_r	M_r	T_r	G_r	M_r
1	744	264	8.25	20 643	2 052	32.1	2 004	872	27.3
2	1 284	502	15.69*	39 234	17 985	281.0*	3 136	2 548	79.6*
4	2 066	46	1.44	60 483	362	5.7	3 724	434	15.6
8	4 086	36	1.13	120 604	6 970	106.1	7 014	50	1.6
16	8 136	396	12.38	234 418	59 982	937.2**	13 978	400	12.5
32	15 876			408 850	4 096	64	27 556		
64				813 604					

b. 分组处理后取平均

r	平方和 T_r (行 2, $n=2^r+4$)					残差平均 \bar{G}_r	均方 M_r
	1—128	2—129	3—130	4—131	5—132		
1	13 578	13 659	13 659	13 663	13 667	7 156	55.9
2	19 952	20 321	20 033	20 325	20 041	6 407	50.1
4	33 820	33 838	33 901	33 842	33 909	4 072	31.8
8	63 282	64 018	63 282	64 022	63 286	24 939	194.8*
16	100 808	100 337	100 816	103 437	103 437	18 709	146.2
32	183 924	187 355	183 924	187 455	187 455	22 684	177.2
64	345 748	354 685	345 748	354 685	345 917	30 122	235.3
128	662 596	677 329	662 765	677 498	662 765		

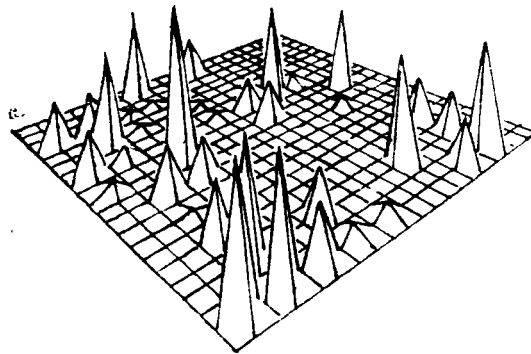


图 4 第一代幼虫分布示意

3 用对分法和样条法判断各种分布的模拟

为了充分说明样条法和对分法判断昆虫种群空间分布型的可行性,用文献中的资料和作者模拟的分布判断各种分布型,并与已有的分布型扩散系数 C 进行对比。

表6中编号1、2、3、4、8、10、17分布为作者模拟的分布,其余编号的分布对应的文献如下:7—[7],12—[8],18,19—[9],16—[10],9—[11],6—[12],5、11、14、15—[3]。

由表6可见,用对分法和样条法对19个分布进行了检验,结果与用扩散系数 C 判断的结果是基本一致的。因此可以认为对分法和样条法判断分布型是可行的。

表6 用对分法和样条法判断各种分布型的模拟对比

编号	$G_r (n=256)$								G_r/G_1	表值	扩散系数 C	相关系数符号	分布型
	1	2	4	8	16	32	64	128					
1	15	16	13	13	6	20	3	17			0.134 4	-	均匀
2	53	45	17	39	11	7	49	9			0.395 0	-	均匀
3	61	58	60	65	54	15	43	28			0.680 0	-	均匀
4	36	30	35	22	30	28	36	25			0.867	-	均匀
5	76	96	91	92	93	44	175 ^①	16	2.215	3.07	0.869 4	-	均匀
6	38	60 ^①	42	54	32	74	36	4	1.579	1.42*	0.940 0	-	聚集 [△]
7	76	70	53	82	130	174 ^①	115	130	2.289	2.40	0.992 5	+	随机
8	19	25	12	14	10	36 ^①			1.890	4.17	1.000 0	+	随机
9	32	33	37	41	27	33	17	81 ^①	2.531	3.84	1.070 0	+	随机
10	72	68	64	76	72	70	65	79			1.095 0	+	随机
11	93	151	105	98	140	106	610 ^①	1 444 ^①	6.559	3.07**	1.200 0		聚集
12	203	198	209	288	267	544 ^①	683	289	3.365	3.07**	1.246 6		聚集
13	111	98	112	148	119	173	213	358 ^①	3.225	2.75*	1.250 0		聚集
14	106	102	119	14	78	58	20	783 ^①	7.387	3.84**	1.430 0		聚集
15	105	162	130	204 ^①	66	499	26	2 704 ^①	1.944	1.83*	1.593 0		聚集
16	52	64	190 ^①	181	110	184	8	0	3.656	1.60**	1.920 0		聚集
17	291	434	460	491	59	173	708	828 ^①	2.845	2.75*	2.110 0		聚集
18	119	147	187	149	323 ^①	89	181	121	2.714	2.00*	2.500 0		聚集
19	131	149	219 ^①	171	185	29	261	25	1.672	1.25*	2.710 0		聚集

①是与 G_1 比较的 G_r 值。*示95%显著,△示判断结果与实际分布不符合。

4 问题讨论

关于对分法和样条法判断昆虫种群空间分布型有以下几个问题需要说明:

(1) 网格大小不同,所得分布型结果不同,如图5所示。格太小或太大均会导致可能错误的结果。格的大小使无虫格占20%~40%为宜,经验结果是平均虫口不低于0.25,且网格不重叠。

(2) 样条法比对分法判断分布更具体,但由于样条选取的随机性,其偶然误差较大,要

克服这个缺点就要多选样条。

(3) 本方法的优点是可以判断聚集范围的大小及位置, 这是分布型指数不具备的。

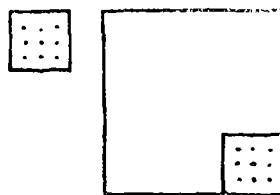


图5 格大小对分布的影响
(左: 均匀分布, 右: 格太大时为聚块)

参 考 文 献

- [1] 南京农学院主编, 1985, 昆虫生态及预测预报, 农业出版社, 141~152.
- [2] 李天生等, 1981, 马尾松毛虫的空间分布型及其在实践上的应用, 林业科学, 4: 343~350
- [3] Ripley, B. D., 1981, Spatial Statistics, John Wiley & Sons.
- [4] Getis, A. et al., 1978, Models of spatial processes, Cambridge University Press, London.
- [5] Mead, R., 1974, A test for spatial pattern at several scales using data from a grid contiguous quadrat, *Biometrics*, 30: 295~307.
- [6] Bartlett, M. S., 1975, Statistical analysis of spatial pattern, Chapman and Hall.
- [7] Gulmon, S. L. et al., 1977, Spatial and temporal relationships between two desert shrubs in Death Valley, California, *J. of Ecology*, 65: 831~838.
- [8] Diggle, P. J. et al., 1983, Some distance-based tests of independence for sparsely-sampled multivariate spatial point patterns, *International Sta. Review*, 51: 11~23.
- [9] Upton, G. et al., 1985, Spatial data analysis by example, John Wiley & Sons Ltd.
- [10] Diggle, P. J. 1979a, On parameter estimation and goodness-of-fit testing for spatial point pattern, *Biometrics* 35: 87~101.
- [11] Keuls, M. et al., 1963, The distance method for estimating densities, *Statistica Neerlandica*, 17: 71~91.
- [12] Kelly, E. P. et al., 1976, A note on Strauss's model for clustering, *Biometrika*, 63: 357~360.
- [13] Zahl, S., 1974, Applications of the S method to the analysis of spatial pattern, *Biometrics*, 30: 513~524.

Application of the Half-half and Spline Methods to Analyse of Insect Population Distribution Patterns

Zhou Guofa

Li Tiansheng

(Branch Campus of Peking University) (The Research Institute of Forestry CAF)

Abstract The half-half and spline methods were applied to analyse the population spatial distribution patterns of the Pine Caterpillar. The two methods can not only be used to test the spatial patterns, but also determine the location and size of clump.

Key words half-half method; spline method; distribution pattern of insect population