

辽宁间日疟残余病例数理概率分布型探讨

辽宁省卫生防疫站 张声阔

辽宁为间日疟流行区。自1965年以来，疫情一直稳定在万分之五以下。一个生产队初发2例病人的，1980年仅有一个。1979~81年，疫源性指数分别为1.08、1.09、1.06。这种高度分散的残余病例的传染关系已很难查清，其分布型式是否改变了传染病分布的聚集性特点而呈现为随机分布，这是疟疾防治工作中一个有理论价值的问题；搞清楚它，对抗疟实践将有指导意义。本文仅就1980年的104例疟疾患者的空间分布，进行普哇松 (Poisson) 和负二项分布 (Negative binomial distribution) 配合适度探讨。

材料和方法

材料来源：各市、地卫生防疫站年末订正的疫情报告。考虑到人口基数的影响，采用以县平均人口数 (61万人) 分组：足61万人即为一个取样单位，不足61万人口者与邻县合并，正负61万×0.5以内者，仍以1个取样单位计算，61万×0.5以上者再计1个取样单位；没有病人的县，将人口总合，以61万除之，为0列的分布频数 (表1)。

研究方法

1. 普哇松分布模式

$$NP_r = Ne^{-\bar{X}} \frac{\bar{X}^r}{r!} \dots\dots\dots (1)$$

$$\sum NP_r = Ne^{-\bar{X}} \left(1 + \bar{X} + \frac{\bar{X}^2}{2!} + \frac{\bar{X}^3}{3!} + \dots + \frac{\bar{X}^r}{r!} \right) \dots\dots\dots (2)$$

其中 NP_r 为理论频数， N 为样本总例数， e 为自然对数之底， $e=2.71828$ ； \bar{X} 为平均数； $r!$ 为 r 的阶乘， $r=0, 1, 2, \dots, N, 0! = 1$

表1 辽宁间日疟残余病例按县平均人口分布

病例数(x)	频数(f)	fx	fx ²
0	18	0	0
1	15	15	15
2	14	28	56
3	3	9	27
4	4	16	64
5	4	20	100
6	0	0	0
7	0	0	0
8	2	16	128
Σ	n = 60	104	390

$$\bar{X} = \frac{\sum fx}{n} = 1.7333$$

$$S^2 = \frac{\sum fx^2 - \frac{(\sum fx)^2}{n}}{n-1} = 3.5548$$

$$S = \sqrt{S^2} = 1.8854$$

2. 负二项分布模式

$$NP_r = \frac{(k+r-1)!}{r! (k-1)!} q^{-k-r} p^r \dots\dots\dots (3)$$

为了简化计算手续，利用前项已求出的理论值连续推导，用(3)式求出 NP_0 的理论值后， NP_1 以后各理论值，可代入下式推导

$$NP_r = \frac{k+r-1}{r} R NP_{r-1} \dots\dots\dots (4)$$

其中 NP_r 及 r 同(1)(2)式， k 是表示聚集程度的参数，求法下述。 R 为常数， $R = p/q$ ，而 $P = \bar{X}/k$ ， $q = p+1$ ； NP_{r-1} 为前项之理论值。

聚集参数 k 的求法有多种^[1]，较粗但多用的为矩法，又称动差法(method of moments)；较精确的是Fisher所创的最大或然法 (Method of maximum likelihood)。此法通过累积频率(A_x)的计算，利用了样本每个组(r)出

注：本文曾在全国疟疾疫情分析会(1981年4月，上海)上交流，此稿根据会上意见做了修改。

现不同频率的全部信息，但计算比较复杂，而且常是利用易于计算的矩法k值开始运算。另有频数法及零频数法等〔1, 2〕。

结果及曲线配合适度检验

一、普哇松分布配合检验

将表1资料按普哇松(2)式展开：

$$NP_0 = Ne^{-\bar{X}} = 60 \times 2.71828^{-1.7333} = 10.60$$

$$NP_1 = Ne^{-\bar{X}} \cdot \bar{X} = 10.60 \times 1.7333 = 18.38$$

$$NP_2 = Ne^{-\bar{X}} \cdot \frac{\bar{X}^2}{2!} = Ne^{-\bar{X}} \cdot \bar{X} \cdot \frac{\bar{X}}{2} = 18.38 \times \frac{1.7333}{2} = 15.93$$

.....

各理论值均记于表2。并以 X^2 测验做配合检验。

二、负二项分布配合检验

k值估计

1. 矩法：从表1提供的 \bar{X} 及 S^2 求k，其公

$$\text{式为 } k = \frac{\bar{X}^2}{S^2 - \bar{X}} = \frac{(1.7333)^2}{3.5548 - 1.7333} = 1.6494$$

其他各参数之计算为

$$q = S^2 / \bar{X} = 3.5548 / 1.7333 = 2.0509$$

$$p = q - 1 = 2.0509 - 1 = 1.0509$$

$$R = p / q = 1.0509 / 2.0509 = 0.5124$$

$$\text{又式 } k = \bar{X} / p = 1.7333 / 1.0509 = 1.6493$$

2. 最大或然法：此法是以试验方法求k，其方程为

$$Z_i = \sum_{x=0}^r \frac{A_x}{ki' + X} - N \ln \left(1 + \frac{\bar{X}}{ki'} \right)$$

式中ln是自然对数，为计算方便，化成常用对数（即以10为底）则

$$Z_i = \sum_{x=0}^r \frac{A_x}{ki' + X} - N \times 2.3026 \times \log \left(1 + \frac{\bar{X}}{ki'} \right) \dots \dots \dots (5)$$

式中 $Z_i = 0$ ， ki' 正是我们要选译的 \hat{K} 值，

一般先以矩法K代入，求出 Z_1 。如果 $Z_1 > 0$ ，应取 $K'_2 > K'_1$ ，如果 $Z_1 < 0$ ，则取 $K'_2 < K'_1$ 。用 K'_2 代入(5)式，再运算一次，得 Z_2 。

$$\text{式中 } A_x = \sum_{i>x}^r f_i$$

A_x 为所有大于x的累积频数和， f_i 为样本中i个个体出现的频数，r为最大的组数。

A _x 之计算			A ₀ = f ₁ + f ₂ + ... + f ₈ = 42
X	f	A _x	
0	18	42	A ₁ = f ₂ + f ₃ + ... + f ₈ = 27
1	15	27
2	14	13	
3	3	10	A ₇ = f ₈ = 2
4	4	6	有了A _x 各值即可代入(5)式。
5	4	2	矩法求出之K值为1.6494，为
6	0	2	计算方便，令 $K'_1 = 1.7$ ，因此
7	0	2	
8	2	0	

$$Z_1 = \frac{4 \cdot 2}{1.7} + \frac{2 \cdot 7}{2.7} + \frac{1 \cdot 3}{3.7} + \frac{1 \cdot 0}{4.7} + \frac{6}{5.7} + \frac{2}{6.7} + \frac{2}{7.7} + \frac{2}{8.7} - 60 \times 2.3026 \times \log \left(1 + \frac{1.7333}{1.7} \right) = 0.02$$

因 $Z_1 > 0$ 故取 $K'_2 > K'_1$

$$\text{令 } K'_2 = 1.9 \text{ 求出之 } Z_2 = -0.32$$

Z_1 为0.02 Z_2 为-0.32，有正有负， $Z = 0$ 必在 K'_1 与 K'_2 之间，此时则可代入线性内插

公式求 \hat{K}

$$\hat{k} = k'_2 + \frac{k'_1 - k'_2}{Z_1 - Z_2} (Z - Z_2) \text{ 其中 } Z = 0$$

$$\hat{k} = 1.9 + \frac{1.7 - 1.9}{0.02 - (-0.32)} \times \{ 0 - (-0.32) \} = 1.7118$$

求出 \hat{K} 值，则可求出其他参数：

$$P = \bar{X} / \hat{K} = 1.7333 / 1.7118 = 1.0126$$

$$q = p + 1 = 2.0126$$

$$R = p / q = 1.0126 / 2.0126 = 0.5031$$

负二项分布则可代入(3)、(4)式求得

$$NP_0 = Nq^{-k} = 60 \times 2.0126^{-1.7118} = 18.12$$

$$NP_1 = NP_0 k R = 18.12 \times 1.7118 \times 0.5031 = 15.61$$

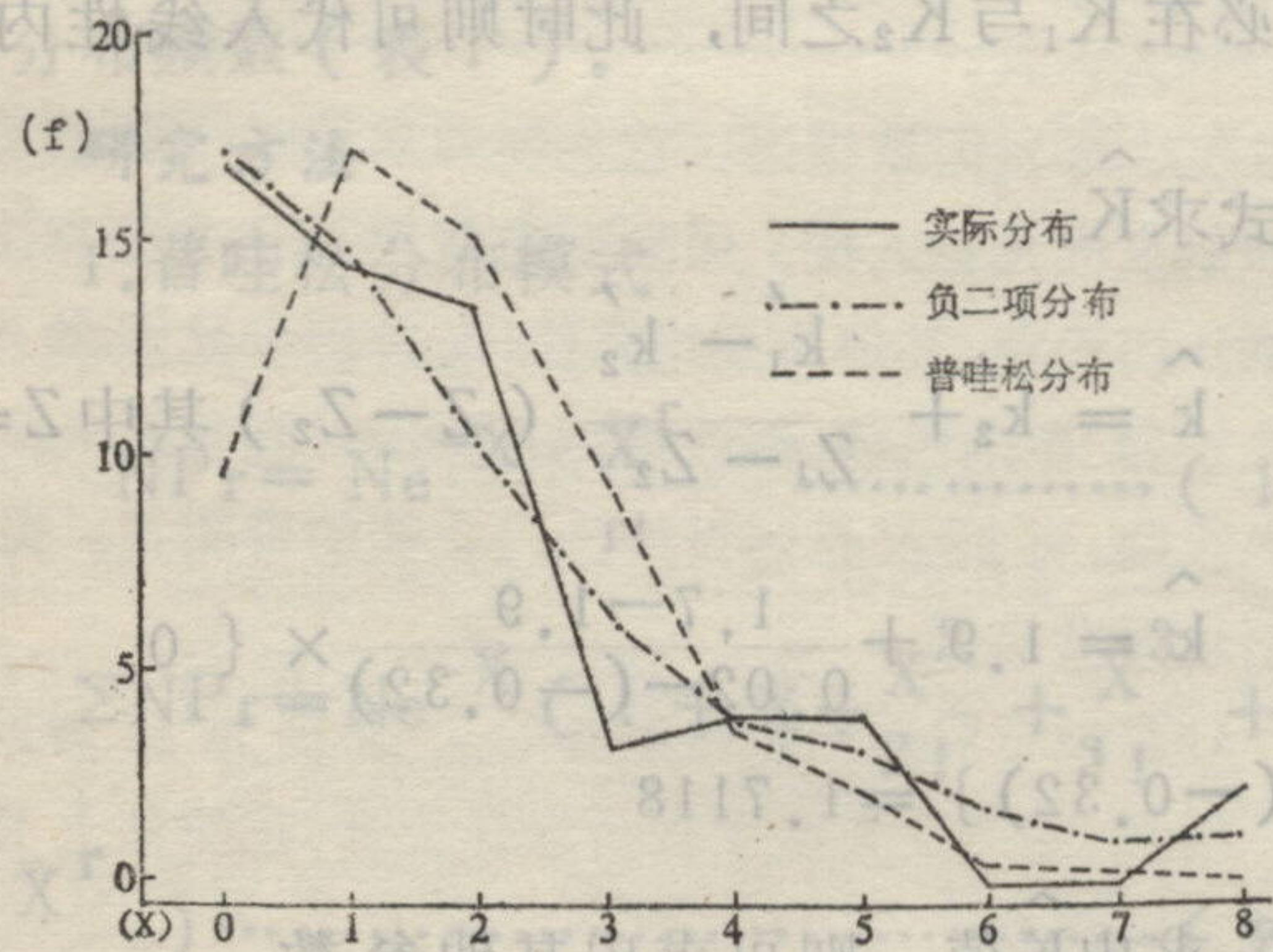
$$NP_2 = NP_1 \frac{K+1}{2} R = 15.61 \times 1.3559 \times 0.5031 = 10.65$$

.....

各项理论值见表2，同时做 χ^2 配合检验：
 经 χ^2 检验：实际分布不符合普哇松分布，符合负二项分布($0.50 > P > 0.25$)见附图。在此计算同时，曾以动差法 $k=1.6494$ 代入(3)、(4)式求理论值，其与实际值之 χ^2 值为3.199，稍大于最大或然法 k 求出之 χ^2 值($2.44 < 3.199$)，但结论一致。

表2 辽宁间日疟残余病例与普哇松及负二项分布之配合

每61万人 中病例数	实际频数 (A)	普哇松分布		负二项分布		
		理论频数(T)	$\frac{(A-T)^2}{T}$	理论频数(T ₁)	$\frac{(A-T_1)^2}{T_1}$	
0	18	10.60	5.17	18.12	0.0008	
1	15	18.38	0.72	15.61	0.0238	
2	14	15.93	0.23	10.65	0.3146	
3	3	9.20	4.18	6.63	1.9875	
4	4	3.99	5.89	3.93	8.99	
5	4	1.38		2.26		
6	0	0.40		1.27		0.1135
7	0	0.10		0.70		
8	2	0.02		0.83		
χ^2			13.17		2.4402	
df			5-2=3		5-3=2	
$\chi^2(0.05)$			7.82		5.99	
p			<0.05		>0.05	
\bar{X}	1.7333					
S	1.8854					
S ²	3.5548					
k				1.7118		



附图 辽宁间日疟病例(1980)与负二项、普哇松分布曲线配合

讨论与小结

普哇松分布是一种机率很少($P \leq 0.01$)而又离散、随机的概率分布。它的特征是每个发生个体在空间任何一点的机率是相等的，并且任一个体的存在决不影响其他个体的分布。辽宁1980年间日疟病例分布虽然给人以随机分布的印象，但经统计检验，0的机率显然高些，它的曲线与普哇松分布模式不相配合，相一致的机率小于5%，这就说明，两者在概率表现上不是一个类型。

负二项分布是一种不随机的聚集性分布，它的特征是真正的核心聚集，即每个抽样单位内一个个体的存在，可增加该单位内其他个体

发生的概率。表示这种聚集程度的参数为 k 值。 k 值越小, 聚集性越大即愈密集, 其随机的可能性就越小; 相反, k 值越大, 其聚集性越差, 随机的可能性就越大; 当 $k \rightarrow \infty$ 时, 其极限则为普哇松分布。我国曾有人研究, 发现钉螺在自然界中^[1], 钩端螺旋体在疫区人群中^[3], 胃癌在某些省区内^[4], 均服从负二项分布, 说明各有其使之聚集的因素, 这就指出了进一步工作的方向。

辽宁省1980年间日症残余病例虽然高度分散, 但已证明它不是普哇松分布, 而与聚集性的负二项分布的理论概率模式相一致, 两者配合满意($0.50 > P > 0.25$)。而且实际分布的统计信息也具有负二项分布的变异数($S^2 = 3.5548$) $>$ 平均数($\bar{X} = 1.7333$)的特点^[1,5]。这就说明, 辽宁1980年间日症病例的分布仍然受到疟疾原发疫区、疫点, 疟疾流行诸因素以及还为人们不甚了解的疟原虫本身的生物学特性所制约, 流

行还没有完全停止, 所以才表现有疫源性疾病的聚集性特点。因此, 在抗疟对策上, 目前仍要坚持事实已证明在辽宁行之有效的“三根治(现患、抗复发、全民)、一预防”的防治方案。至于传播机制问题, 在有长潜伏期的情况下, 少数病例或带虫者传播后隔年发病^[6], 正是其正常表现。

(本文在全国疟疾疫情分析会上承蒙顾杏元副教授提出宝贵意见, 谨致谢意)

参 考 文 献

1. 苏德隆: 中华卫生杂志, 8(1): 59, 1963.
2. 丁岩钦: 昆虫种群数学生态学原理与应用, 90页, 科学出版社, 1980.
3. 乔树民、陈书兴: 《卫生防疫资料》, 1页, 辽宁省卫生防疫站, 1963.
4. 龚惠心: 负二项分布在肿瘤流行病学中的应用, 内部资料, 1980.
5. 耿贯一主编: 流行病学, 下册, 336页, 人民卫生出版社, 1980.
6. 辽宁省卫生防疫站: 中华预防医学杂志, 15(3): 140, 1981.

西藏地区慢性布病患者免疫水平的初步观察

北京同仁医院 左大鵬

笔者对西藏当雄县部分慢性布氏杆菌病患者的非特异性免疫进行了初步观察。用PHA皮试法检测100例正常人和62例慢性布病患者的非特异性细胞免疫水平, 结果表明布病组阳性率明显低于正常对照组(表1)。进一步对62例慢性患者做不同临床分型的对比观察, 两组间的阳性率无显著差异(表2)。

表1 布病组和正常对照组PHA皮试阳性率比较

	例数	阳性数	阳性率%	P值
布病组	62	9	14.52	P < 0.01
正常对照组	100	43	43.00	

表2 不同临床分型患者PHA皮试阳性率比较

	例数	阳性数	阳性率%	P值
慢性活动型	23	0	0	P > 0.05
慢性稳定型	39	9	23.08	

用检测免疫球蛋白的方法检测117例慢性布氏杆菌病患者的非特异性体液免疫水平, 结果分别为: IgG 11.4973 ± 4.7736 毫克/毫升, IgM 0.8965 ± 0.3722 毫克/毫升, IgA 1.5838 ± 0.6290 毫克/毫升, 均(为 $\bar{X} \pm S.D$) 与10例当地正常人的结果对比, 无显著差异($P > 0.05$)。117例慢性患者中, 慢性活动型50例, 慢性稳定性67例, 两组间IgG水平对比, 也无显著差异($P > 0.05$)。观察表明, 慢性布氏杆菌病患者的非特异性体液免疫水平无显著变化。为了观察病程与免疫水平之间的关系, 将117例慢性患者按病程长短分为一年以上, 三年以上, 五年以上, 十年以上四个组, 分别进行IgG和布氏杆菌血清凝集效价(试管法)的对比观察, 发现随病程迁延, IgG水平逐渐下降, 十年以上患者甚至低于正常水平, 四个组之间对比有一定差异($P < 0.05$); 而血清学效价虽然也随病程迁延逐渐下降, 但四个组之间无显著差异($P > 0.05$)。