中华流行病学杂志1992年第13卷第1期

系列讲座

VII. 二项分布及其应用

上海医科大学* 詹绍康

二项分布是一种重要的统计概率分布，在属性资料统计分析中占有重要地位。

二项分布的性质

从阳性比例为π的总体中随机抽取含量为n的样本，其中出现阳性个体数为r的概率是

$$P_r=C_n^r(1-\pi)^{n-r}\pi^r \qquad (1)$$

由于P_r正好是二项式$[(1-\pi)+\pi]^n$展开后的各项，因此称r的概率分布为二项分布（binomial distribution）。

〔例1〕 某地中学生近视眼患病率为π=0.30，随机抽取n=4名学生，其中近视眼患者数r可以是0，1，2，3，4五种情况，这五种情况的概率理论上可用式（1）来计算：

$r=0,\ P_0=C_4^0(1-0.3)^{4-0}0.3^0=0.7^4=0.2401$

$r=1,\ P_1=C_4^1(1-0.3)^{4-1}0.3^1=4\times0.7^3\times0.3=0.4116$

$r=2,\ P_2=C_4^2(1-0.3)^{4-2}0.3^2=6\times0.7^2\times0.3^2=0.2646$

$r=3,\ P_3=C_4^3(1-0.3)^{4-3}0.3^3=4\times0.7\times0.3^3=0.0756$

$r=4,\ P_4=C_4^4(1-0.3)^{4-4}0.3^4=0.3^4=0.0081$

上述5个概率，正好是二项式$(0.7+0.3)^4$的展开式的各项。

一、二项分布的三个基本假定：上述阳性数r服从二项分布的条件是：

1.每个个体只有阳性和阴性两种可能结果；

2.每个个体呈现阳性的概率相等；

3.这n个个体中一个个体是否阳性与其它个体无关。

二、二项分布的参数及其性质：

1.二项分布有两个参数，一个是总体阳性比例π，另一个是样本含量n。π的取值在0和1之间，n是正整数。π和n决定后，二项分布也就决定了。

2.二项分布变量r的均数是nπ，标准差是$\sqrt{n\pi(1-\pi)}$

3.若以二项分布变量r除以样本含量n，得样本比例p=r/n。p与r的概率分布相对应，样本比例p的均数是π，标准差是$\sqrt{\pi(1-\pi)/n}$。

4.π越接近0.5，n越大，则反复随机抽样后样本比例p的分布越接近对称，有统计学者提出，当π为不同数值时，只要n足够大，可近似地把样本比例p看作正态分布（表1）。

表1 正态近似要求的样本含量

π	n
0.5	30
0.4	50
0.3	80
0.2	200
0.1	600
0.05	1400
~0*	nπ=80

*如果总体比例π极小，n为无穷大的时候，阳性数r服从普哇松分布

按二项分布原理估计总体比例

在实际工作中，人们常需要从样本比例p=r/n来估计总体比例π。总体比例的可信区间的下限π_L和上限π_U，可按二项分布原理求得，统计学家已把π_L和π_U值制成了表，使用相当方便。我们选其中一小部分列表2。

〔例2〕 在某小学随机抽查12名学生，其中4名为近视眼患者，问全校学生近视眼患病率是多少？

本例中可以n=12，r=4从表2查得10～65，即认为全校学生近视眼比例在10%到65%之间，统计学上称10%～65%为总体比例的95%可信区间。

*邮政编码 200032

表2　　总体比例的95%可信区间

n \ r	0	1	2	3	4	5	6	7	8	9	10
1	0～98										
2	0～84	1～99									
3	0～71	1～91	9～99								
4	0～60	1～81	7～93								
5	0～52	1～72	5～85	15～95							
6	0～46	0～64	4～78	12～88							
7	0～41	0～58	4～71	10～82	18～90						
8	0～37	0～53	3～65	9～76	16～84						
9	0～34	0～48	3～60	7～70	14～79	21～86					
10	0～31	0～45	3～56	7～65	12～74	19～81					
11	0～28	0～41	2～52	6～61	11～69	17～77	23～83				
12	0～26	0～38	2～48	5～57	10～65	15～72	21～79				
13	0～25	0～36	2～45	5～54	9～61	14～68	19～75	25～81			
14	0～23	0～34	2～43	5～51	8～58	13～65	18～71	23～77			
15	0～22	0～32	2～41	4～48	8～55	12～62	16～68	21～73	27～79		
16	0～21	0～30	2～38	4～46	7～52	11～59	15～65	20～70	25～75		
17	0～20	0～29	2～36	4～43	7～50	10～56	14～62	18～67	23～72	28～77	
18	0～19	0～27	1～35	4～41	6～48	10～54	13～59	17～64	22～69	26～74	
19	0～18	0～26	1～33	3～40	6～46	9～51	13～57	16～62	20～67	24～71	29～76
20	0～17	0～25	1～32	3～38	6～44	9～49	12～54	15～59	19～64	23～69	27～73

又如，某县随机抽查10名农民，其中8名有蛔虫感染，要估计全县农民的蛔虫感染率也可查表2。用n=10及r=8无法查表2，我可以把不感染蛔虫者人数作为r，即r=2。以n=10，r=2从表2查得3～56，即全县农民中不感染蛔虫者的比例估计为3%～56%，也就是总体感染率估计为44%～97%。

对样本比例p与某一已知总体比例π_0的差别作统计检验

〔例3〕 某地过去都对锡克氏反应阳性8～15岁学生皮下接种0.1ml吸附精制白喉类毒素，免疫后十天锡克氏试验阴转率49.5%，今研制一种新制剂，试验10人，十天阴转8人，问用新制剂能否提高10天阴转率。

本例中，49.5%是已知的总体比例，即$\pi_0=0.495$，样本含量n=10，阴转数r=8，样本比例p=8/10=0.8，用统计学术语来说，就是要对样本比例p与总体比例π_0的差别作统计检验。检验的原理就是计算在比例为π_0的总体中随机抽样获得样本比例与总体比例之差达0.8−0.495=0.305及比0.305更大的概率。如果此概率很小（如小于0.05），则认为p与π_0的差别有统计意义；如果此概率不小（如大于0.05），则认为p与π_0的差别无统计意义。出现前一种结果时，认为新制剂10天阴转率高于0.495，出现后一种结果时，认为用新制剂10天阴转率仍为0.495。

样本含量为10时，与总体比例$\pi_0=0.495$之差$\geqslant$0.305的样本比例是$P_8=8/10=0.8$，$P_9=9/10=0.9$，$P_{10}=10/10=1.0$，以公式（1），可算得

$$P_8=C_{10}^{8}(1-0.495)^{10-8}\,0.495^{8}=0.0414$$

$$P_9=C_{10}^{9}(1-0.495)^{10-9}\,0.495^{9}=0.0090$$

$$P_{10}=C_{10}^{10}(1-0.495)^{10-10}\,0.495^{10}=0.0009$$

$$P=P_8+P_9+P_{10}=0.0414+0.0090+0.0009=0.0513$$

因$P>0.05$，在0.05水准下，认为p与π_0的差别无统计意义，没有理由说用新制剂的10天阴转率不再是0.495。

又如某乡过去大量调查得钉螺的血吸虫感染率为1.2%，今在该乡随机捕得钉螺240枚，查得感染螺1枚，问现在钉螺的血吸虫感染率是否比过去下降了？

据题意，$\pi_0=0.012$，n=240，r=1，p=1/240=0.00417。检验的假设是H_0：$\pi=\pi_0$，即假设目

前的钉螺血吸虫感染率仍为0.012。本例中 $p-\pi_0=-0.00783$，与π_0差别达-0.00783（或写作$p-\pi_0\leqslant-0.00783$）的样本只有两种，$p_1=1/240=0.00417$及$p_0=0/240=0$。从比例为$\pi_0=0.012$的总体中随机抽样得这两种样本的概率可用式（1）来计算：

$$P_1=C_{240}^{1}(1-0.012)^{239}0.012=0.16081$$

$$P_1=C_{240}^{0}(1-0.012)^{240}0.012^{0}=0.05517$$

$$P=P_1+P_0=0.16081+0.05517=0.21598$$

显然$P>0.05$，p与π_0的差别无统计意义，不拒绝H_0，没有理由说该乡钉螺感染率已下降。

二项分布的正态近似

如果π和n符合表1所列条件，从比例为π的总体里随机抽取含量为n的样本，样本比例 p近似于均数为 π 标准差为$\sigma_p=\sqrt{\frac{\pi(1-\pi)}{n}}$ 的正态分布，σ_p的估计值为s_p，$s_p=\sqrt{\frac{p(1-p)}{n}}$

利用二项分布的正态近似性，在医学统计资料分析中可有如下几方面的应用：

一、用正态近似法对总体比例π作区间估计。总体比例π的$100(1-\alpha)\%$可信区间为$p\pm u_\alpha s_p$，若定α为0.05，则π的95%可信区间为$p\pm u_{0.05}s_p=p\pm1.96s_p$；若定α为0.01，则π的99%可信区间为$p\pm u_{0.01}s_p=p\pm2.58s_p$。

〔例4〕某医师随机抽查142名百日咳患者，其中临床症状轻微的非典型患者有44人，问全部百日咳患者中非典型患者的比例可能是多少？

本例若用95%可信区间来估计总体比例π，则$\alpha=0.05$，$p=44/142=0.31$，$s_p=\sqrt{0.31(1-0.31)/142}=0.0388$，$u_{0.05}=1.96$，$p\pm u_{0.05}s_p=0.31\pm1.96\times0.0388=0.23$，0.39。意即全部百日咳患者中非典型患者比例估计在23%～39%之间。

二、用正态近似法对样本比例p与某个已知总体比例π_0的差别作统计检验，可用统计量u

$$u=\frac{p-\pi_0}{\sigma_p} \qquad (2)$$

此u统计量近似于标准正态分变量，即$u_{0.05}=1.96$，$u_{0.01}=2.58$。

〔例5〕某地几年来狂犬病患者中40%是16岁内的青少年，今年随机抽查120名狂犬病患者，有63名不满16岁，问今年16岁内患者的比例是否增高了。

按题意，假设$\pi_0=0.4$，$n=120$，$r=63$，$p=63/120=0.525$，$\sigma_p=\sqrt{\pi_0(1-\pi_0)/n}=\sqrt{0.4\times0.6/120}=0.0447$，故有

$$u=\frac{p-\pi_0}{\sigma_p}=\frac{0.525-0.40}{0.0447}=2.80$$

本例若用双侧检验，$u_{0.05}=1.96$，$u_{0.01}=2.58$，$u>u_{0.01}$，$P<0.01$，p与π_0（$\pi_0=0.4$）的差别有统计意义，拒绝H_0，认为今年狂犬病患者中青少年的比例比0.40已有所上升。

又如，据以往大量资料统计，狂犬病患者中男女之比为3：1，今年随机抽查150名狂犬病患者，其中45名为女性，问该病患者中男女性比例是否仍为3：1？

据题意，假设男女性患者比例为3：1，即女性患者比例为25%，$\pi_0=0.25$，$n=150$，$r=45$，$p=45/150=0.30$。

$\sigma_p=\sqrt{0.25(1-0.25)/150}=0.0354$。故有

$$u=\frac{p-\pi_0}{\sigma_p}=\frac{0.30-0.25}{0.0354}=1.41$$

本例若作双侧检验，$u<u_{0.05}$，$P>0.05$，不拒绝假设，没有理由说狂犬病患者男女性之比不是3：1。

三、用正态近似法对两个样本比例p_1与p_2的差别作统计检验，可用公式

$$u=\frac{p-1p_2}{S(p_1-p_2)} \qquad (3)$$

式中

$$s(p_1-p_2)=\sqrt{PQ\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

$$P=\frac{n_1p_1+n_2p_2}{n_1+n_2},\qquad Q=1-P$$

u统计量仍近似于标准正态分布变量。

〔例6〕随机抽查甲乙两地传染病新病例140和160例，其中菌痢新病例数分别为50和32，问甲乙两地传染病新病例中菌痢所占比例是否相同？

据题意，本例所作假设是H_0：$\pi_1=\pi_2$，即假设两地传染病新病例中菌痢所占比例相同。同时有$p_1=50/140=0.3571$，$p_2=32/160=0.2000$，$P=\frac{50+32}{140+160}=0.2733$，$Q=1-0.2733=0.7267$，$s(p_1-p_2)=\sqrt{0.2733(1-0.2733)\left(\frac{1}{140}+\frac{1}{160}\right)}=0.0516$，故有

$$u=\frac{p_1-p_2}{s(p_1-p_2)}=\frac{0.3571-0.2000}{0.0516}=3.04$$

本例作双侧检验$u_{0.01}=2.58$，$u>u_{0.01}$，$P<0.01$，两样本比例差别有统计意义，拒绝H_0，认为甲 乙 两地传染病新病例中菌痢所占比例是不相等的。

混合样品分析

在某些场合的调查中，样品阳性率很低，检测 的量很大，研究者希望能以较少的人力、物力和 时间来获得检测结果，本节所讲的混合样品分析 法，就是可供选择的一种方法。

应用混合样品分析法的条件是：把几份原始样 品混合成一份（混合样品），只要几份样品中有1份（或多于1份）阳性，这几份原始样品组成的混 合 样 品就呈阳性。

〔例7〕 某地采集了10 000份原始样品，目的是了解原始样品的阳性率。从专业知识已知，即使 把10份原始样品混合成1份，只要其中有1份（或多于1份）阳性，该混合样品就呈阳性。因此，把每10份原始 样品混合成1份，10 000份原始样品变成了1000份 混 合 样品。对这1000份混合样品作检测结果，64份阳 性，即混合样本阳性率$p_c=\frac{64}{1000}=0.064$。此时，原始样品 阳性率p可用下列公式计算：

$$(1-p)^n=1-p_c$$

或

$$p=1-\sqrt[n]{(1-p_c)}$$

本例中 $n=10$，$p_c=0.064$，故

$$p=1-\sqrt[10]{1-0.064}=0.00659=6.59‰$$

原始样品阳性率估计为6.59‰。

如果研究者还打算了解哪几份原始样品阳 性，那么在开始时就应把10 000份原始样品的每一份分成两半，一半用于组成混合样品，另一半备用。当 检测到64份混合样品阳性时，就把这64份混合样品所含的640份原始样品备用的另一半取出作检测，这样便可检出 这10 000份原始样品中的全部阳性样品。本来，对10000份原始样品要作10 000次检测，采用 混 合 样 品 分 析法，两批共检测了1000＋640＝1640（次），省了 8360次检测所需的人力、物力和时 间。当 然，作 混 合 样品分析时，每一份原始样品量要多一点，并能分成 两半，备用的一半要能保存且不影响检测结果。

二项分布的配合

二项分布是一种随机事件的概率分布。因 此，在某些情况下，可通过某事件是否近似于二项分布的 分析，来判断其分布的随机法。在医学上，造成疾 病分布不随机性的因素可能是疾病的传染性、遗传 性、共同生活条件等。

二项分布的配合，是分析和判断某事件（疾 病）是否属二项分布的一种方法。

假定某研究者调查了一批五口之家，记录了其 中近视人数，得如下资料（表3），问近视眼 在 各 户之分布是否随机分布。

表3　　有不同数量近视患者的户数

近视患者数	0	1	2	3	4	5	合计
户　　数	6	8	9	5	6	7	41

对表3资料作二项分布配合的步骤如下：

1.估计近视眼患病率。本例中调查总人数 为41×5＝205人，近视眼患病总人数为1×8＋2×9＋3×5＋4×6＋5×7＝100人，所以估计的近视眼患 病率为$\pi=100/205=0.4878$。

2.计算n＝5，p＝0.4878，q＝0.5122时，r＝0，1，2，3，4，5的概率，所用公式是：

$$P_r=C_n^r q^{n-r} p^r$$

对本例

$$P_0=C_5^0\,0.5122^5\ 0.4878^0=0.0353$$

$$P_1=C_5^1\,0.5122^4\ 0.4878^1=0.1679$$

$$P_2=C_5^2\,0.5122^3\ 0.4878^2=0.3197$$

$$P_3=C_5^3\,0.5122^2\ 0.4878^3=0.3045$$

$$P_4=C_5^4\,0.5122^1\ 0.4878^4=0.1450$$

$$P_5=C_5^5\,0.5122^0\ 0.4878^5=0.0276$$

3.计算有不同数量近视患者的理论户数Tr

$$T_0=0.0353\times41=1.4$$

$$T_1=0.1679\times41=6.9$$

$$T_2=0.3197\times41=13.1$$

$$T_3=0.3045\times41=12.5$$

$$T_4=0.1450\times41=5.9$$

$$T_5=0.0276\times41=1.1$$

把上述各项计算结果与表3原始资料一起可列成表4，表中r表示近视患者数，Ar表示观察到的 户 数，Pr表示按n＝5，p＝0.4878的二项分布计算的理论 概 率，Tr表示按上述二项分布概率计算的理论户数。

4.用χ^2检验来分析配合二项分布的合适性：

$$\chi^2=\sum\frac{(A-T)^2}{T}$$

表4　　配合二项分布用计算表

r	Ar	Pr	Tr
0	6	0.0353	1.4
1	8	0.1679	6.9
2	9	0.3197	13.1
3	5	0.3045	12.5
4	6	0.1450	5.9
5	7	0.0276	1.1

χ^2检验中要求理论（频）数不小于1，如果小于1，应作合理的合并。这种χ^2检验中，如果患病率p是从样本估计的，该χ^2统计量服从自由度为k－2的χ^2分布；如果患病率不是从样本估计的，而是某个假设的数值或已知的数值，那么自由度为k－1，其中k是最后用于计算χ^2统计量的观察数或理论数个数。本例中未作任何合并，k=6，故自由度$\gamma=k-2=6-2=4$。

$$\chi^2=\frac{(6-1.4)^2}{1.4}+\frac{(8-6.9)^2}{6.9}+\frac{(9-13.1)^2}{13.1}+$$

$$\frac{(5-12.5)^2}{12.5}+\frac{(6-5.9)^2}{5.9}+\frac{(7-1.1)^2}{1.1}=52.72$$

$\chi^2_{0.01(4)}=13.28$　$\chi^2>\chi^2_{0.01}$，　$P<0.01$，认为近视患者在各户的分布并非二项分布。

家族聚集性分析

上面所作的二项分布之配合中，规定n=5，即只能调查五口之家，这在实际工作中是比较困难的，因为家庭人口数是有不同的。本节介绍的方法，就是对具有不同n的二项分布作配合的家族聚集性分析。

〔例8〕　某研究者在某地随机调查了一批家庭，清点了每个家庭中的近视眼患者数，得如下资料(表5)。

表5中右边的一列是有不同数量近视眼患者的户数；没有近视患者的有30户，1例近视患者的有61户，……。底下一列是具有不同数量人口的户数：1口之家8户，2口之家30户，……。表内没有括号的数字是调查（观察）到的户数，8户1口之家中，4户没有近视患者，4户各有一例近视患者；30户2口之家中，10户没有近视患者，13户各有1例近视患者，7户各有2例近视患者，……。

对本例资料分析步骤如下：

1.估计近视眼患病率p。本项调查的总人数为：

表5　　近视眼家族聚集性分析用表

近视眼患者数 r	不同家庭人口(n)的户数 1	2	3	4	5	6	7	合计
0	4 (3.97)	10 (7.39)	3 (4.28)	4 (2.67)	6 (1.23)	3 (0.51)	0 (0.17)	30 (20.22)
1	4 (4.03)	13 (15.00)	15 (13.02)	10 (10.83)	8 (6.26)	9 (3.09)	2 (1.21)	61 (53.44)
2		7 (7.61)	10 (13.22)	8 (16.50)	9 (12.72)	9 (7.85)	6 (3.69)	49 (61.59)
3			7 (4.48)	14 (11.17)	5 (12.91)	1 (10.63)	4 (6.24)	31 (45.43)
4				8 (2.83)	6 (6.55)	6 (8.09)	4 (6.33)	24 (23.80)
5					7 (1.33)	3 (3.28)	1 (3.86)	11 (8.47)
6						3 (0.55)	0 (1.31)	3 (1.86)
7							6 (0.19)	6 (0.19)
合计	8	30	35	44	41	34	23	215

●括号内的数字文内有说明

1×8+2×30+3×35+4×44 +5×41+6×34+7×23=919（人），其中近视患者数为1×61+2×49+3×31+4×24+5×11+6×3+7×6=463（人）。因此估计近视眼患病率为p=463/919=0.5038，估计的近视眼不患率为q=1−p=1−0.5038=0.4962。

2.以p=0.5038，q=0.4962,n=1,2，……7分别计算r=0，1，……的（二项分布）概率。列于表6。

表6　　p=0.5038 时 的 二 项 分 布 概 率

r	n						
	1	2	3	4	5	6	7
0	0.4962	0.2462	0.1222	0.0606	0.0301	0.0149	0.0074
1	0.5038	0.5000	0.3721	0.2462	0.1527	0.0909	0.0526
2		0.2538	0.3778	0.3750	0.3101	0.2308	0.1603
3			0.1279	0.2538	0.3148	0.3125	0.2713
4				0.0644	0.1598	0.2379	0.2755
5					0.0325	0.0966	0.1678
6						0.0164	0.0568
7							0.0083
合计	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

表6中各项概率的计算公式是：

$$P_r = C_n^r \; q^{n-r} \; p^r$$

当n=1，

$P_0 = C_1^0 \cdot 0.4962^1 \cdot 0.5038^0 = 0.4962$

$P_1 = C_1^1 \cdot 0.4962^0 \cdot 0.5038^1 = 0.5038$

当n=2，

$P_0 = C_2^0 \cdot 0.4962^2 \cdot 0.5038^0 = 1 \cdot 0.4962^2 = 0.2462$

$P_1 = C_2^1 \cdot 0.04962^1 \cdot 0.5038^1 = 2 \cdot 0.4962 \cdot 0.5038 = 0.5000$

$P_2 = C_2^2 \cdot 0.04962^0 \cdot 0.5038^2 = 1 \cdot 0.5038^2 = 0.2538$

当n=3，

$P_0 = C_3^0 \cdot 0.04962^3 \cdot 0.5038^0 = 1 \cdot 0.4962^3 = 0.1222$

$P_1 = C_3^1 \cdot 0.4962^2 \cdot 0.5038^1 = 3 \cdot 0.4962^2 \cdot 0.5038^1 = 0.3721$

$P_2 = C_3^2 \cdot 0.4962^1 \cdot 0.5038^2 = 3 \cdot 0.4962^1 \cdot 0.5038^2 = 0.3778$

$P_3 = C_3^3 \cdot 0.4962^0 \cdot 0.5038^3 = 1 \cdot 0.5038^3 = 0.1279$

其余依此类推。

3.计算理论户数，即按二项分布理论，p=0.4962时，家庭人口数n=1，2，3，4，5，6，7时，近视患者数为r的户数理论上是多少。本例只要用表6中的概率分别乘以表5底行不同家庭人口的户数即可，结果列于表5括号内。

当n=1，

8×0.4962=3.97

8×0.5038=4.03

当n=2，

30×0.2462=7.39

30×0.5000=15.00

30×0.2538=7.61

其余依此类推。

4.横向计算行合计。即调查得0例患者的有30户，按二项分布理论算得0例患者的应有20.22户；调查得1例患者的有61户，按二项分布理论算得1例患者的应有53.44户……，前者称实际数A，后者称理论数T。

5.作χ^2检验以判断调查所得资料是否符合二项分布。所用公式是：

$$\chi^2 = \sum \frac{(A-T)^2}{T}$$

此χ^2统计量服从自由度为k-2的χ^2分布，k为最后用于计算χ^2的理论数（或实际数）个数。用此公式时，要求理论数不小于1，否则应先作合理的合并。本例中r=7时，理论数为0.19，已小于1，应作合并，（1.86+0.19）=2.05，相应的实际数也应合并，（3+6）=9，因此本例的k不再是8,而是7,如果患病率p（本例中的近视率0.5038）是从调查的样本中算得的，χ^2检验的自由度为k−2；如果患病率p是某个已知或假设的值,自由度为k−1。本例自由度为7−2=5。

$$\chi^2=\frac{(30-20.22)^2}{20.22}+\frac{(61-53.44)^2}{53.44}+\frac{(49-61.59)^2}{61.59}+\frac{(31-45.43)^2}{45.43}$$

$$+\frac{(24-23.80)^2}{23.80}+\frac{(11-8.47)^2}{8.47}+\frac{(9-2.05)^2}{2.05}=37.28$$

因$\chi^2_{0.01(5)}=15.09$，$\chi^2>\chi^2_{0.01(5)}$，$P<0.01$。认为调查所得资料并不服从二项分布，所以认为近视眼患者在各户的分布并不随机。可能由于某种原因，有些家庭的成员易患近视眼，另一些家庭的成员不易患近视眼。一些家庭比另一些家庭的成员更易患近视眼的原因可能是遗传因素、共同生活环境等，对于某些疾病，可能是其传染性。

内蒙呼盟地区STD流行病学调查分析

内蒙古自治区呼盟流行病防治研究所*

那　森　郭　莉　陈　方　王作林　王福山　沙　丽　那　林

为掌握呼盟地区性传播疾病（STD）的流行动态，以利制订防治措施提供依据，于1990年在该盟所辖12个旗市范围进行STD流行病学调查，其结果报告如下。

一、对象与方法：

1.对象：特殊职业人群（汽车驾驶员、旅饭店等公共场所服务人员）；特殊人群（在押犯人、被收容审查人员）及其他人群。

2.方法：被调查对象在指定的地点时间内受检，首先填写登记表，而后临床检查。检查和诊断方法按卫生部防疫司编《性病防治手册》一书进行，

二、结果：本次在呼盟12个旗市共检查36 277人，检出STD患者234例，发病率0.65%。

1.发病情况：

①各旗市发病情况：扎兰屯市发病率1.29%、莫旗0.65%、鄂伦春旗0.303%、额左旗0.52%、牙克石市0.32%、鄂温克旗0.89%、陈旗0.59%、满洲里市0.13%、额右旗0.04%、西旗0.21%、东旗0.13%、海拉尔市发病数47例、不详9例。

②病种分布：梅毒8例（构成比3.42%），淋病93例（39.74%），尖锐湿疣74例（31.62%），NGU 1例（0.43%），其它58例（24.79%）。

2.流行病学特征：

①地区分布：城市发病率0.83%，农区0.65%，牧区0.59%，林区0.35%，不详3例（$\chi^2=26.69$，$P<0.005$）。

②各年龄组发病情况：19岁以下发病率0.71/万、20～29岁0.88/万、30～39岁0.54/万、40～49岁0.25/万、50～59岁0.30/万、60岁以上4.41/万。

③与文化程度的关系：高中文化程度的发病率.37%、初中0.30%、小学2.69%（$\chi^2=25.71$，$P<0.005$）。

④与婚姻关系：已婚者发病率0.59%，未婚者发病率1.02%（$\chi^2=1.16$，$P>0.05$）。

⑤与性别的关系：男性发病率0.55%，女性为3.16%（$\chi^2=13.98$，$P<0.005$）。

⑥与职业关系：无业人群发病率为33.33%，工人25.66%、个体18.18%、农民13.73%、干部7.36%、供销5.88%、服务4.05%、司机0.28%。

三、讨论分析：在呼盟所属12个旗市范围共检查36 277人，检出STD患者234例，发病率0.65%，说明该地区普遍流行STD。流行的病种除艾滋病无条件检查外其余数病均有流行，其中淋病占首位（发病专率0.26%）；其次为尖锐湿疣（0.203%）；梅毒占第三位（0.022%）；NGU目前最低（0.003%）。

流行病学特征①有显著的地区差别，城市＞农区＞牧区＞林区；②发病年龄集中在20～39岁范围；③与文化程度、职业有很大关系，但与婚姻无关系；④女性发病率与男性发病率有极显著性差异，两者之比为6：1。

* 海拉尔市，邮政编码 021008