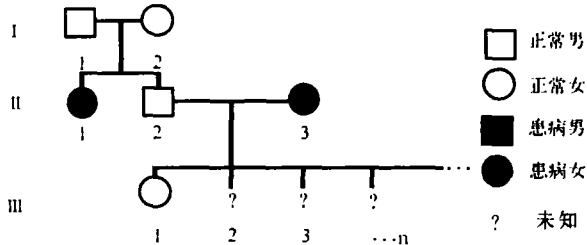


系谱分析中概率问题的贝叶斯分析一例

何风华 (华南师范大学生命科学学院 广东广州 510631)

《生物学通报》2007年第2期“遗传概率的一个认识误区”一文,给出了一个家系概率问题的答案,但没有详尽的分析过程,不利于初学者的学习掌握,本文对此予以补充和完善。

问题:下面是某家族单基因遗传病系谱图,基因用A、a表示。请据图回答:Ⅱ2号的基因型概率是多少?



系谱分析表明该病的遗传方式是常染色体隐性遗传。如果这个家系的第Ⅲ代表型未知,可以根据孟德尔定律得出Ⅱ2的基因型为AA的概率是1/3,为Aa的概率是2/3。若Ⅱ2和Ⅱ3生了患病个体,那么可以确定Ⅱ2就是致病基因携带者,此时Ⅱ2再生育1个患儿的概率为1/2。本例已知第Ⅲ代有n个表型正常的子女,但这一信息不能帮助我们准确推定Ⅱ2的基因型概率。很显然随着健康孩子的连续出生,Ⅱ2为致病基因携带者的风险越来越小,但不能就此断言他不是携带者。系谱概率分析中的基本法则是加法定理和乘法定理。这个家系中,在已知Ⅱ2有n个正常子女的情况下,估计其基因型概率还需要用到贝叶斯(Bayes)定理,分析过程通常称为贝叶斯分析。

贝叶斯定理表示假设某一事件E已经发生,该事件可能由一组原因事件(C_i)引发,那么在事件E已经发生的情况下某一特定原因 C_i 的概率 $P(C_i|E)$ 等于未发生事件E时该原因出现的概率 $P(C_i)$,乘以该事件对于该原因的概率 $P(E|C_i)$,除以事件E发生的总概率。事件E发生的总概率等于所有引起事件E发生的潜在因素的概率乘以事件E对于每个因素的概率的总和。贝叶斯定理可写为等式:

$$P(C_i|E) = \frac{P(C_i)P(E|C_i)}{\sum P(C_j)P(E|C_j)}$$

贝叶斯公式也叫做逆概率公式,因为它实际上是在知道了结果的情况下推断原因事件的概率,它反映了试验之后对原因发生可能性大小的新认识。贝叶斯定理用于医学遗传学中,E常指某一疾病的受累或未受累个体,或有某实验结果的人, C_i 表示特定的基因型。

贝叶斯公式用于该家系, C_i 指Ⅱ2的基因型,E指Ⅱ2生育n个健康子女这一事件。贝叶斯分析通常包括4个层次的概率。在未知Ⅱ2子女表型的情况下,我们认为Ⅱ2基因型的概率是前概率。Ⅱ2是致病基因

携带者Aa的前概率是2/3;如果Ⅱ2的基因型就是Aa,那么他生育1个正常子女的概率为1/2,这个在假设条件下产生的特定表型概率称为条件概率。根据独立事件概率的乘法定理,Ⅱ2有2个正常子女的条件概率是 $1/2 \times 1/2$,有n个正常子女的条件概率是 $(1/2)^n$ 。整合这个家系提供的3个世代的信息,则Ⅱ2是致病基因携带者,同时他有n个正常子女的概率为前概率与条件概率的乘积,或 $2/3 \times (1/2)^n$ 。前概率和条件概率说明的2个事件同时发生的概率叫做联合概率。同理,Ⅱ2是纯合体AA的前概率是1/3;如果Ⅱ2就是正常的纯合体AA,那么他的后代必然全部表型正常,其条件概率为1,Ⅱ2是纯合子AA和有n个正常子女的联合概率为 $1/3 \times 1 = 1/3$ 。根据互斥事件的加法定理,Ⅱ2表型正常(基因型为AA或Aa),同时他有n个正常子女的概率为Ⅱ2各基因型的联合概率之和,或 $2/3 \times (1/2)^n + 1/3$ 。再根据贝叶斯定理,系谱中Ⅱ2在已生育n个健康子女的情况下是隐性基因携带者的后概率等于Ⅱ2为Aa的联合概率除以Ⅱ2各基因型的联合概率之和,或 $\frac{2/3 \times (1/2)^n}{2/3 \times (1/2)^n + 1/3} = \frac{1}{1+2^{n-1}}$;Ⅱ2已有n个健康子女,并且Ⅱ2为正常纯合子AA的后概率为 $\frac{1/3}{2/3 \times (1/2)^n + 1/3} = \frac{2^{n-1}}{1+2^{n-1}}$ 可以通过下表的方法比较直观地推导出Ⅱ2基因型的后概率。

	Ⅱ2基因型概率的 Bayes 分析	
	Ⅱ2是致病基因携带者 Aa	Ⅱ2是纯合子 AA
前概率 $P(C_i)$	2/3	1/3
条件概率 $P(E C_i)$	$(1/2)^n$	1
联合概率 $P(E, C_i)$	$2/3 \times (1/2)^n$	1/3
后概率 $P(C_i E)$	$1/(1+2^{n-1})$	$2^{n-1}/(1+2^{n-1})$

如果Ⅱ2已生育2个健康的子女,那么他是致病基因携带者的概率由2/3降到了 $1/(1+2^{2-1})=1/3$ 。假若他再想要1个孩子,这个孩子患病的概率是多少呢?由于纯合子aa与携带者婚后生出隐性患儿的风险为1/2,故Ⅱ2再生育患儿的风险为 $1/3 \times 1/2 = 1/6$ 。可见,再生育子女的患病风险大大降低。当双亲或双亲之一的基因型不明确时,无论遗传病的遗传方式怎样,估计双亲生育某种患儿的概率要用到贝叶斯分析。贝叶斯定理在医学遗传学中有十分广泛的用途,有兴趣的读者可继续参阅有关专业书籍。

主要参考文献

- 1 董莱哈特 T.D.,柯林斯 F.S.,金斯伯格 D..医学遗传学原理.孙开来等译.北京:科学出版社,2001:276—280.
- 2 李巍,何蕴韶.遗传咨询.河南:郑州大学出版社,2003:57—76.
(E-mail:hefenghua2000@163.com) (BH)