



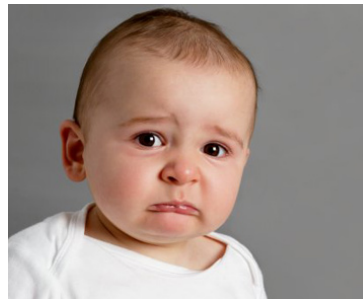
Sally Clark: 一个被错误的统计证据毁掉一生的母亲*

2007年3月16日，Sally Clark的家人悲痛地宣布，Sally于当日清晨在家中去世。Sally Clark，这位可怜的母亲在被精神问题困扰多年后，因酒精中毒离开人世。这一切悲剧始于一项对她错误的审判：被控谋杀了两个儿子，之后虽然被证明是无辜的，但是她也始终没能从中恢复过来。下面，我们详述事情的经过，并从统计的角度进行讨论。

1996年，Sally与Steve的第一个儿子，Christopher，在出生11周后突然去世。去世的原因被认定为自然死亡，医生，邻居对他们一家丧子之痛同情万分。然而，3年后不幸再次降临，她仅8周大的第二个儿子，Harry，又以同样的方式猝死。这次，他们没有得到同情，与此相反，却受到了质疑。4周后，这对夫妇被警方逮捕，并被控谋杀两个儿子。最后，Sally Clark被判处终身监禁。饱受丧子之痛的母亲在监狱中承受生理和心理的双重煎熬，虽几经上诉，直到2003年，新的证据才证明她的清白，随后她被释放。巨大的伤痛使这位母亲的生活异常艰难，只能靠酒精麻醉缓解痛苦。2007年，Sally Clark因酒精中毒离开人世，从此彻底摆脱进入天堂见到了两个儿子。



Sally Clark



妈妈是被冤枉的。。。。。

记者Geoffrey Wansell把Sally Clark案件称为“现代英国法律史上最大的不公”。这是因为在审判中，儿科专家Roy Meadow¹错误地使用了统计证据和统计推理，因而误导了陪审团，最终导致悲剧发生。

不合理的统计证据——独立性问题。Roy Meadow在证据中有如下陈述：在一个家庭中，一个婴儿突然夭折是一场悲剧，两个婴儿猝死值得怀疑，三个婴儿的死亡则可认定是谋杀，除非能够证明事实并非如此。²他声称，像Clark这样一个富裕、不抽烟的家庭，一个婴儿猝死的概率仅为1/8543，两个婴儿猝死的概率为 $(1/8543)^2 = 1/73000000$ 。Roy Meadow假设两个婴儿的死亡是独立事件，也就是说，第一个婴儿死亡的概率不会影响第二个婴儿死亡的概率，这如同扔硬币，第一次得正的概率不会影响第二次得正的概率。伦敦每年大约有700000婴儿出生，根据上述结果，Meadow认为一个家庭中两个婴儿均猝死的事件每100年才发生一次。

*本文作者高磊、宋培培和杨晓康。

¹Roy Meadow的[详细资料](#)。

²这一说法又称为[Meadow定律](#)。



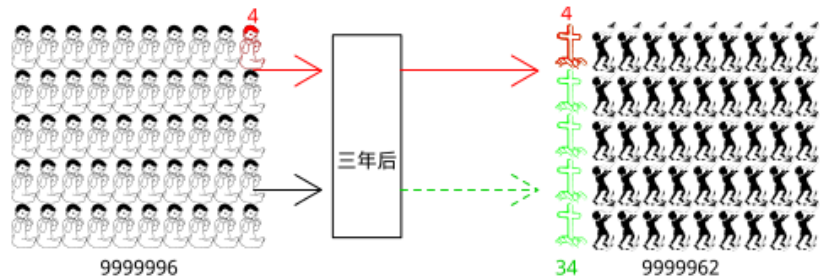
Roy Meadow 提供的统计证据忽略了两个婴儿死亡之间的关联性，把两个婴儿死亡过于简单地当做独立事件。事实上，当一个家庭中出现了婴儿猝死事件，就有理由认为这个家庭存在着引起婴儿死亡的高风险因素（基因、环境等因素），因此再次发生婴儿死亡的概率将有所提高，仍然使用 1/8543 这个概率是愚蠢的。这比愚蠢还糟糕，这是坏科学。基于以上考虑，Roy Meadow 得出的骇人听闻的概率 1/73000000 就不可信，并且所谓的 100 年发生一次也是一种凭空想象。然而，当这些证据呈现在法庭上时，没有人提出质疑。



提供的统计证据是错误的

错误的统计推理——检察官谬误。³1/73000000 意味着什么？英国一位著名记者报道，Roy Meadow 认为 1/73000000 是 Sally 无罪的概率。不必责备 Roy Meadow，也无需苛责陪审团，我们很多人也认为 Roy Meadow 的推理是正确的，甚至有人感觉这是朴素的概率论常识。然而，这是亟需纠正的错误，用一张图示来解释其中的缘由。

我们假设，在某段时间内有 1000 万名婴儿出生。需要注意的是，为了叙述方便，我们用一名婴儿代替两名婴儿，当我们讲到一名婴儿怎么样时，其实是说两名婴儿的情况。比如，我们说，一名婴儿的母亲具有谋杀动机，其实我们想说的是，生育两个婴儿的母亲想杀死自己的孩子。



1000 万婴儿出生后 3 年内的存活状况

把婴儿们分成两组，分组的标准是他们的母亲是‘好’（S）还是‘坏’（N），‘好’指母亲疼爱自己的孩子，‘坏’指母亲具有谋杀自己孩子的动机。众所周知，虎毒不食子，几乎没有母亲会谋杀自己的孩子。用概率的语言表述，就是‘坏’母亲概率极低，CONI（Care of Next Infant）的研究表明，拥有‘坏’母亲的婴儿比例为

$$P(N) = 0.0000004.$$

换句话说，在 1000 万名婴儿中，平均而言，只有 4 名婴儿的母亲具有谋杀动机，用左图右上角的红色小人表示这些不幸的婴儿。对于手无缚鸡之力的婴儿来说，如果他们的母亲想要谋害他们，几乎不费吹灰之力。因此，如果一名婴儿的母亲想要害他，这名婴儿是不可能存活的，三年后，这些

³法院在审判过程中犯了统计错误，被称为“检察官谬误”。检察官谬误的详细解释请参阅 [Prosecutor's fallacy](#)。



婴儿几乎都被其母亲害死，我们用右图左上角的红色坟墓表示这 4 名被母亲谋杀的婴儿。因此，一名拥有‘坏’母亲的婴儿三年内必然死亡，用概率表示：

$$P(D|N) = 1.$$

相对来说，另外 9999996 名婴儿无疑是幸福的。生活水平也许有好有坏，但他们有一个疼爱自己的母亲，把这些婴儿用左图的黑色小人表示。因此，拥有好母亲的婴儿比例：

$$P(S) = 1 - P(N) = 0.9999996 \approx 1.$$

然而，我们知道，对于新生婴儿来说，有许多不确定性因素会导致他们的死亡，甚至最终也查不出原因所在，儿科医生干脆把不明原因的死亡称为婴儿猝死综合征（SIDS）。CONI（Care of Next Infant）的研究表明，婴儿死于 SIDS 的概率为：

$$P(D|S) = 0.0000034.$$

这也就意味着，在 9999996 名婴儿中，平均而言，三年后有 34 名婴儿死于 SIDS，我们用右图左侧绿色坟墓表示 34 名死于 SIDS 的婴儿。除此之外的 9999962 名婴儿，将会进入儿童时期，用右图右侧掷飞机的小人表示。

上图全景式地展现了婴儿出生后三年内的存活状况，而且可以回答我们在 Sally Clark 案件中遇到的问题：如果观察到一名婴儿死亡了，那么他死于 SIDS 还是母亲谋杀呢？在上图右侧 38 名死亡婴儿中，34 名死于 SIDS，只有 4 名死于母亲谋杀，比例分别为

$$P(S|D) = \frac{34}{38} = 0.89,$$

$$P(N|D) = \frac{4}{38} = 0.11.$$

换句话说，当观察到一名婴儿死亡时，我们会做出如下判断：这名婴儿死于 SIDS 的概率为 89%，而死于谋杀的可能性为 11%。

我们还可以根据概率运算法则——贝叶斯公式得到与上述结论一致的结果。利用贝叶斯公式分别计算 $P(S|D)$ 和 $P(N|D)$ ：

$$\begin{aligned} P(S|D) &= \frac{P(D|S) * P(S)}{P(D|S) * P(S) + P(D|N) * P(N)} \\ &= \frac{0.0000034 * 0.9999996}{0.0000034 * 0.9999996 + 1 * 0.0000004} \\ &= 0.89; \end{aligned}$$

$$\begin{aligned} P(N|D) &= \frac{P(D|N) * P(N)}{P(D|S) * P(S) + P(D|N) * P(N)} \\ &= \frac{1 * 0.0000004}{0.0000034 * 0.9999996 + 1 * 0.0000004} \\ &= 0.11. \end{aligned}$$

以上过程解释了 Sally Clark 有更大的无罪概率，但仍未讲明 Roy Meadow 以及陪审团犯错的原因。我们继续探究这一问题。事实上，陪审团成员并没有犯‘错’，他们使用了频率学派（统计有两派：频率学派和贝叶斯学派）的方法进行估计，更确切地说，他们使用极大似然估计来寻找一名婴儿死亡的原因。⁴

⁴关于极大似然估计与贝叶斯估计的区别，请参阅神文[数学之美番外篇：平凡而又神奇的贝叶斯方法](#)。



极大似然估计建立在似然函数基础之上，意在寻找使似然函数最大的那个参数。在 Sally Clark 案件中，就是寻找使

$$P(D|\text{原因})$$

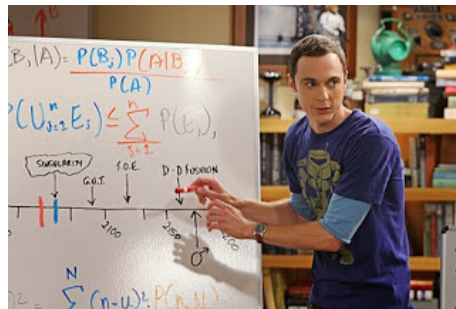
最大的那个原因。前面已经讨论过，如果原因是“谋杀” (N)，那么似然函数值为

$$P(D|N) = 1;$$

如果是 SIDS(S)，那么似然函数值是

$$P(D|S) = 0.0000034.$$

显然，当原因是“谋杀”时，似然函数值更大。所以，我们会倾向于认为 Sally Clark 谋杀了自己的孩子。



Sheldon 用贝叶斯公式估计自己的寿命

极大似然估计的致命缺点是没有考虑“原因”本身可能性的大小，当这些原因的可能性有显著性差别时（本例中，具有谋杀动机的母亲和关爱孩子的母亲比例差距颇大），极大似然估计的结论就会大错特错，本例中陪审团的判断是其生动体现。与之相反，贝叶斯估计方法综合考虑了原因本身可能性大小（“先验”）和似然函数值（“似然”），会给出一个更加可信的估计和判断。

参考资料：

- [Beyond reasonable doubt](#)
- [Multiple sudden infant deaths – coincidence or beyond coincidence?](#)
- [维基百科：Sally Clark](#)
- [Sally Clark 的主页](#)