
手把手教你整理出一份合格的生存分析数据

作者：张华 赵一鸣 来源：临床流行病学和循证医学

本文原地址：<https://www.iikx.com/news/statistics/1667.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

在统计咨询时，发现研究者的生存分析数据多数不太规范，很多研究者不知道怎么整理生存分析的数据，从而导致无法进行分析。生存分析的数据看似很难整理，其实只要抓住规律，整理起来并不困难。

生存分析中我们不仅关心是否发生结局，还会关心结局发生的时间，是否发生结局及时间这两个数据共同组成了生存分析的因变量(Y)。自变量与其它分析方法的数据格式相同，在此不作讨论。

对于结局，分析软件也只认两种情况：发生和未发生。一般把生发结局事件标为1，未发生结局事件标为0;未发生结局事件的数据又称为删失数据或截尾数据。

对于时间，当发生结局时，时间一列应是结局发生的时间，而未发生结局时，时间就是最后一次随访成功的时间。

上面是整理数据的原则，但具体处理时还会有很多情况，今天从几个病例的角度给大家讨论一下这个问题。

现有一个队列研究，以某疾病的复发或死亡作为终点事件，3个月随访一次，随访期为1年，有如下几个病例：

病例号	3个月	6个月	9个月	12个月
01	良好	复发		
02	良好	良好	复发	死亡
03	良好	良好	车祸死亡	
04	良好	良好	良好	良好
05	良好	良好	未成功	
06	未成功			

病例01比较好理解，在第6月随访时复发了，此时结局应为“1”，时间应为6月；

病例02出现复发和死亡，结局应为“1”，时间怎么办呢？这时应为第一次出现事件的时间，即9月；

病例03在9月时因车祸死亡，这里并不是因某病的死亡，当病人因与本疾病无关的原因复发或死亡时，结局状态应记为“0”，时间为9月；

病例04在12月时仍未复发或死亡，因此结局记为“0”，时间为12月；

病例05在9个月时未随访成功，不知道其9个月时的状态，此时应记录其最后一次随访成功时的状态，即结局为“0”，时间为6月；

病例06则一次未能随访成功，只能记录开始入组时的状态，即结局为“0”，时间为0月。

上述数据整理完成后即：

病例号	结局(end)	时间 (time) 单位: 月
01	1	6
02	1	9
03	0	9
04	0	12
05	0	6
06	0	0

后面的两列数据即可以作为因变量进行单因素和多因素生存分析。



更多 统计方法 请访问 <https://www.iikx.com/news/statistics/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](http://iikx.com)转发