生存曲线是怎么画出来的

作者：陶立元 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/1778.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！

生存分析在我们的临床研究中经常会被使用。在做生存分析时不仅考虑了是否“ 生存 ”，同时还考虑了“ 生存时间 ”的长短。正是由于这广为流传的名字，初学者对生存分析才容易产生误解，以为生存分析只可以用来分析生存资料，即生死大事。其实不然!

确切地说**生存分析**应该叫**事件时间分析**(Time-to-Event Analysis)。何谓事件时间分析，就是不光分析事件是否发生，还分析事件发生的时间间隔长短。比如我们要分析A、B两种措施治疗网球肘的疗效，以上肢恢复正常活动为指标，我们收集了所有患者治疗后是否恢复正常活动，以及何时恢复的正常活动这两个信息。这个研究里面的事件就是：上肢恢复正常活动，时间就是:治疗距离恢复的时间间隔，时间单位根据课题需要自行设定。如果你测量的够准确可以以秒为单位，哈哈，是不是有点儿过分了!

此时有人会说，那我可以把这个数据拆成两个分析：首先分析事件是否发生，使用卡方检验;其次分析事件发生时间长短时，使用计量资料分析方法(t检验或非参数)。这是不对的，因为上述的任何一个分析，都忽略了另一个的存在，没有综合考虑，而且此类数据往往存在删失的情况。

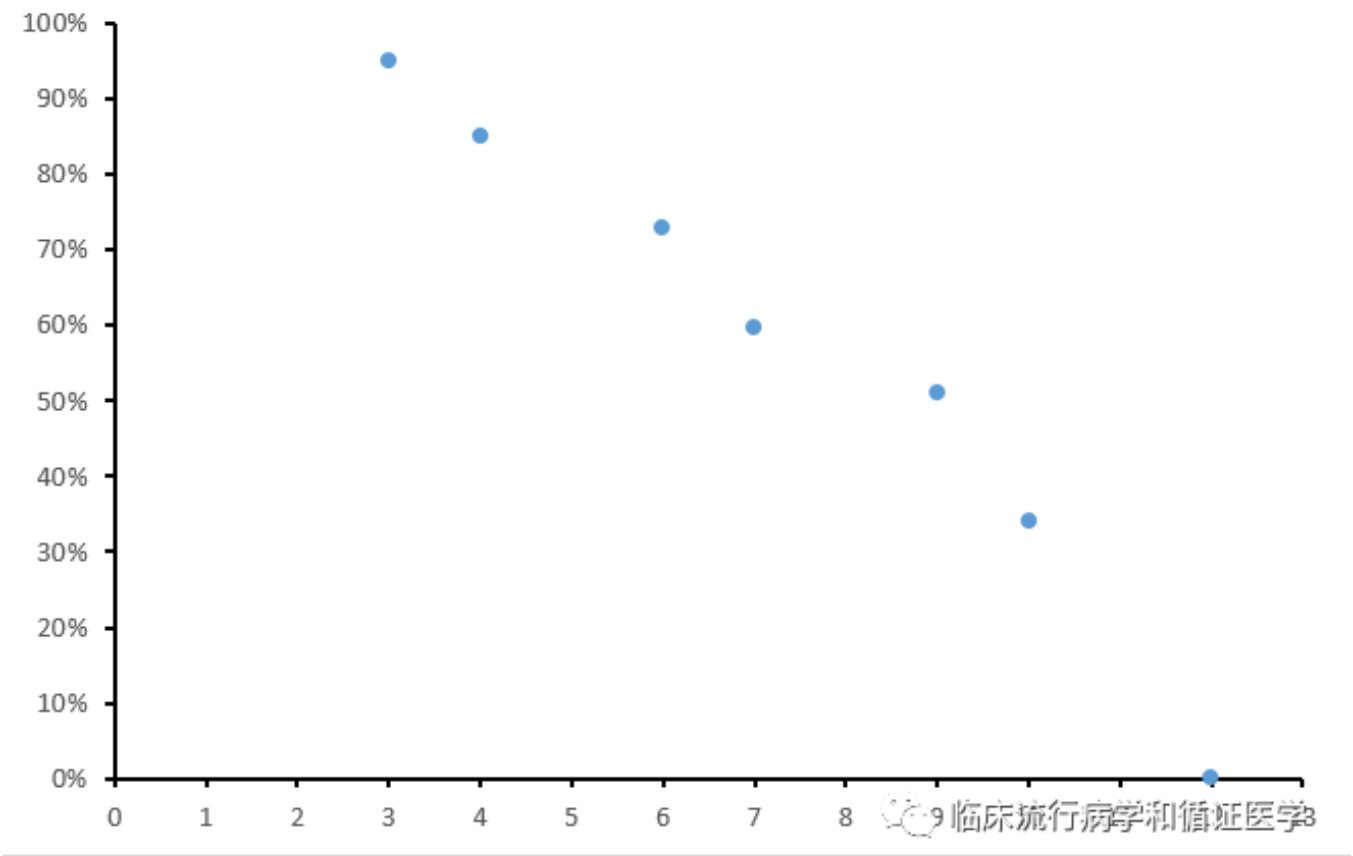
何谓**删失**
(Censoring)，就是没有观察到事件发生的，这些病例都叫删失病例。删失病例有两种：1是到达随访结束时，这个病人也没有出现终点事件;2是在随访过程中病人没有出现终点事件，但是后来就联系不上该病人了，无法继续进行随访直至随访结束。下面假设一组网球肘治疗后的数据，数据是假设。

病人编号	是否恢复	恢复时长	期初未恢复人数	期末未恢复人数	未恢复时点概率	未恢复累计概率
1	1	3 周	20	19	0.950	0.950
2	0	4 周				
3	1	4 周				
4	1	4 周	19	17	0.895	0.850
5	0	5 周				
6	0	5 周	16	16	1.000	0.850
7	0	6 周				
8	1	6 周				
9	1	6 周	14	12	0.857	0.729
10	1	7 周				
11	1	7 周	11	9	0.818	0.596
12	0	8 周				
13	0	8 周	9	9	1.000	0.596
14	1	9 周	7	6	0.857	0.511
15	1	10 周				
16	1	10 周	6	4	0.667	0.341
17	0	11 周				
18	0	11 周	4	4	1.000	0.341
19	1	12 周				
20	1	12 周	2	0	0.000	0.000

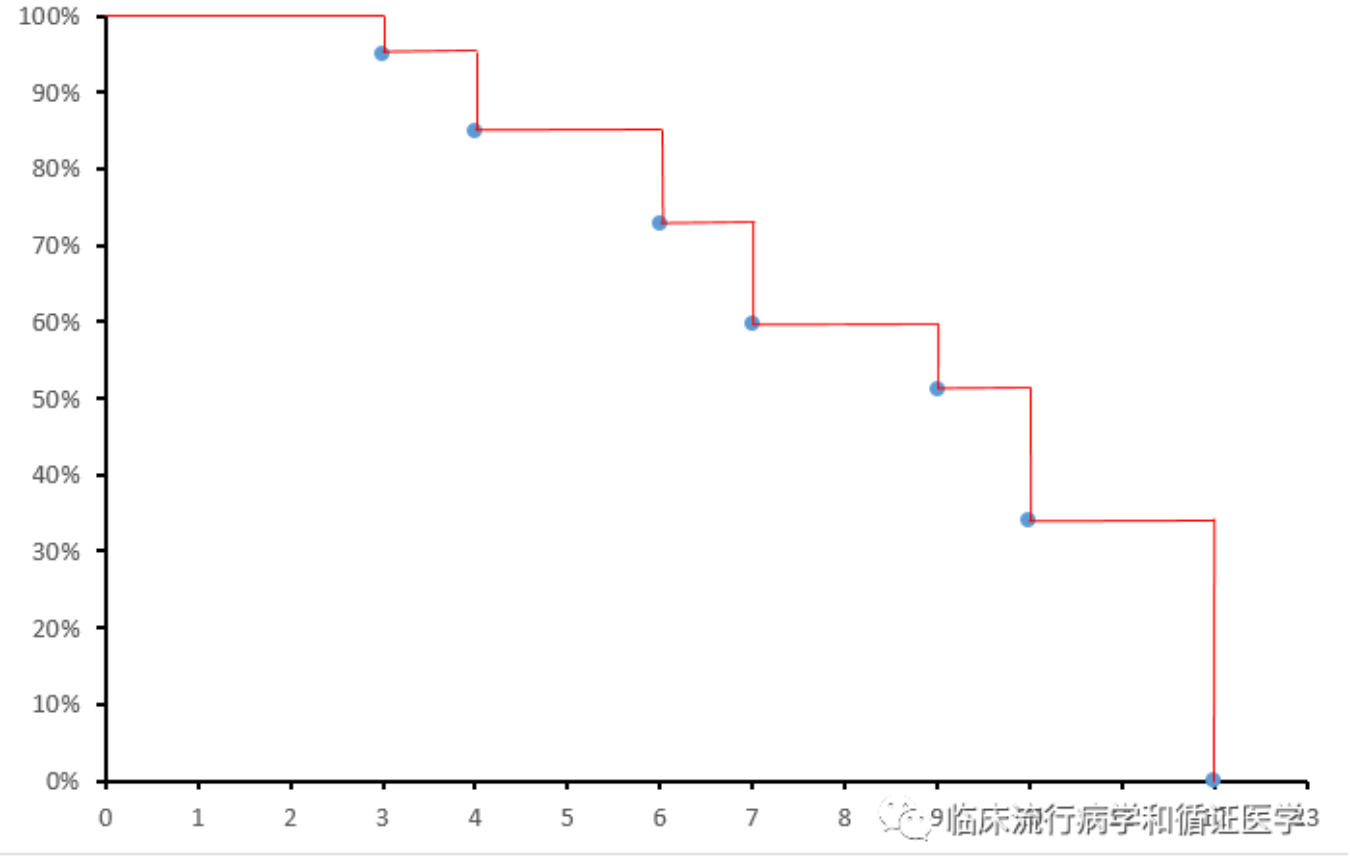
临床流行病学和循证医学

在上述的表中有时点概率和累积概率的区别。时点概率就是在特定的周内，期末有多少人除以该周内期初时的总人数。累积概率就是时点概率的依次乘积，如6周时的累积概率=1周的时点概率×2周的时点概率×3周的时点概率×4周的时点概率×5周的时点概率×6周的时点概率=5周的累积概率×6周的时点概率。

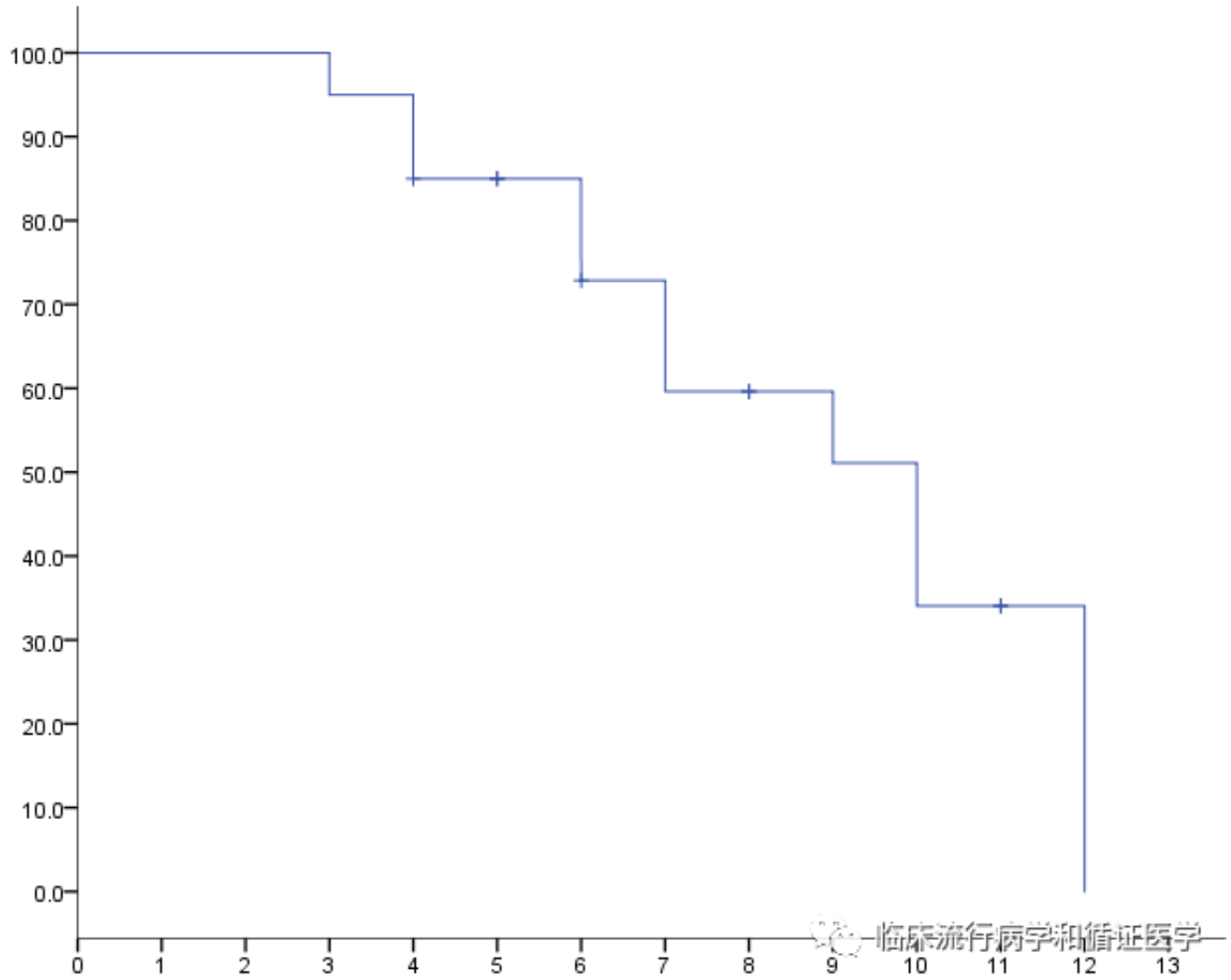
算出了上表的累积概率，我们就可以以时间为x轴，以累积概率为y轴，做如下散点图。



将上述散点图用线连接如下，便是生存曲线。当然这条曲线里面没有删失值的表达。



下面这副图是SPSS制作的同样数据的生存曲线，+表示删失数据。



从上面简单的生存曲线的制作过程，想告诉大家生存曲线里累积生存率的重要概念。同时各位在使用生存分析时，还容易出现以下两种错误：

1、特定时点累积生存率的计算

临床研究中常用的有3年生存率和5年生存率等。很多人会直接用该时点还存活的人/总人数。如本例中6周未恢复率=6周时未恢复人数/总人数=11/20=55.0%。而上表中6周未恢复率=72.9%。

两者相差还是挺多的，主要原因是由于删失数据的存在，55.0%的算法未考虑删失数据的影响。如果没有删失数据，两者的结果应该是一样的。同样存在问题的是两组间特定生存率的组间比较，如两组患者3年生存率的比较。此时需要将随访时点设定到3年再使用软件进行比较。

2、中位生存时间的计算。

中位生存时间是指50%人还生存时的时长。有两种做法，一种是在上述曲线上在y轴的50%位置画一条横线，看其与生存曲线交点在x轴的数值。另一种叫线性内切法，就是按照等比例的原则进行计算。

如上述示例，计算中位未恢复时间t。由上表可知9周时未恢复率为0.511，10周时未恢复率为0.341，计算公式如下：

$$\frac{(9-10)}{9-t}=\frac{0.511-0.341}{0.511-0.5}$$

临床流行病学和循证医学

计算得t=9.06周。

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转发