
如何借助ROC曲线筛选最优界值

作者：王晓晓，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/1668.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！

都说可以借助ROC分析的曲线坐标确定最优界值(optimal cut off)

，但看着SPSS的结果，有些小伙伴不免沉思，到底如何着手呀?(至于如何得到曲线坐标，在ROC分析界面选中ROC曲线的坐标点即可)

一、顾此失彼可取吗，这个还得看您

首先，输出结果中曲线坐标一共有三列，第一列为界值，即大于或等于该值时判定为阳性，并据此和金标准结果比较，得出相应的灵敏度和特异度，曲线坐标第二列和第三列分别给出灵敏度和1-特异度(灵敏度同敏感度，特异度同特异性)

是不是瞬间豁然开朗了?根据咱们对灵敏度和特异度的要求，选择最优界值就可以了呀。比如该病病死率很高、严重威胁人类健康和生命，这时候漏诊的危害相当大，那我们就需要尽可能提高灵敏度、减少漏诊，也就是要求第二列数值比较大。我们看到如果临界值在19-41之间时，灵敏度大于等于0.9，可以在19-41之间选择。

有一些情况，比如该病尚未有确证有效的治疗方案，且现有治疗花费巨大时，我们需要尽可能减少误诊，尽可能提高特异度，也就是要求1-特异度比较小，即第三列数值比较小。

二、不想偏颇，就想齐头并进，可行吗

但是有些小伙伴不想顾此失彼，同时要求高灵敏度和高特异度，这时候又当如何寻找最优界值?除了大家最熟悉的约登指数(Youden ' s index)，再给大家推荐另外两种易操作的方法：product index和Euclidean ' s index。

首先，约登指数大家比较熟悉。约登指数=灵敏度+特异度-1，约登指数最大时对应的界值距离机会线($y=x$)较远，灵敏度和特异度均较高，常作为筛选最优界值的根据。

其次，product index等于灵敏度和特异度的乘积，该指数最大时对应的界值距离机会线较远，灵敏度和特异度也均较高，

此外，还有一个不知道中文名的指标，Euclidean ' s index等于 $(1-\text{灵敏度})^2 + (1-\text{特异度})^2$ 的平方，该指数最小时对应的界值距ROC曲线左上角最近，灵敏度和特异度也较高。

小伙伴肯定又有问题了，这三种方法又该选哪个呢?目前，因为缺乏评价标准，所以，三种指标孰优孰劣不可知。当然，以后，有可能通过一定的统计方法比较不同指标的优劣，现阶段小编建议大家选其中一个即可。

三、对理论过敏，有软件操作吗

最后，小编还是分享一下自己平常筛选最优界值的流程，有点简朴。

第一步，在SPSS输出界面复制曲线坐标。

IBM SPSS Statistics 查看器

文件(E) 编辑(E) 查看(V) 数据(D) 转换(T) 插入(I) 格式(O) 分析(A) 直销(M) 图形(G) 实用程序(P) 扩展(X) 窗口(W) 帮助(H)

输出

- 日志
- ROC 曲线
 - 标题
 - 备注
 - 个案处理摘要
 - ROC 曲线
 - 曲线下方的区域
 - 曲线的坐标

曲线的坐标

检验结果变量: score

大于或等于此 值时为正 ^a	敏感度	1 - 特异性
19.00	1.000	1.000
22.50	1.000	.950
26.50	1.000	.900
28.50	1.000	.850
29.50	1.000	.800
31.00	1.000	.750
33.00	1.000	.700
34.50	1.000	
36.00	1.000	
37.50	.950	
39.00	.900	
41.00	.900	
42.50	.850	
43.50	.800	
44.50	.750	
45.50	.700	
46.50	.650	.650

剪切
复制
选择性复制...
粘贴到后面
创建/编辑自动脚本...
样式输出(E)...
导出...
编辑内容(O)

IBM SPSS Statistics 处理程序就绪 Unicode:ON H: 19.61, W: 6.56 cm

第二步，粘贴到excel中，首先，删掉前三行和后两行有文字的内容，然后插入标题行：cut off、se、1-sp、Youden ' s index、Euclidean ' s index、product index。

通过公式编辑和填充即可知道不同界值对应的三个index指标。比如，通过Euclidean ' s index筛选最优界值，将Euclidean ' s index排序后，Euclidean ' s index最小时对应的即可作为最优界值。也可以分别将Youden ' s index、product index排序后，取其最大时对应的界值。

工作簿1 - Excel						
文件	开始	插入	页面布局	公式	数据	审阅
视图	登录					
A1	曲线的坐标					
1	曲线的坐标					
2	检验结果变量:					
3	大于或等于此值时 为正 ^a	敏感度	1 - 特异性			
4	19.00	1.000	1.000			
5	22.50	1.000	0.950			
6	26.50	1.000	0.900			
7	28.50	1.000	0.850			
8	29.50	1.000	0.800			
9	31.00	1.000	0.750			
10	33.00	1.000	0.700			
11	34.50	1.000	0.600			
12	36.00	1.000	0.500			
13	37.50	0.950	0.400			
14	39.00	0.900	0.350			
15	41.00	0.900	0.300			
16	42.50	0.850	0.300			
17	43.50	0.800	0.250			
18	44.50	0.750	0.250			
19	45.50	0.700	0.100			
20	46.50	0.650	0.050			
21	47.50	0.600	0.050			
22	51.50	0.550	0.050			
23	56.00	0.500	0.000			
Sheet1						
就绪 计数: 6 100%						

工作簿1 - Excel							
文件 开始 插入 页面布局 公式 数据 审阅 视图 登录							
SUM		✕ ✓ f_x		=B2+(1-C2)-1			
	A	B	C	D	E	F	G
1	cut off	se	1-sp	Youden' s index	Euclidean' s index	product index	
2	19.00	1.000	1.000	=B2+(1-C2)-1			
3	22.50	1.000	0.950				
4	26.50	1.000	0.900				
5	28.50	1.000	0.850				
6	29.50	1.000	0.800				
7	31.00	1.000	0.750				
8	33.00	1.000	0.700				
9	34.50	1.000	0.600				
10	36.00	1.000	0.500				
11	37.50	0.950	0.400				
12	39.00	0.900	0.350				
13	41.00	0.900	0.300				
14	42.50	0.850	0.300				
15	43.50	0.800	0.250				
16	44.50	0.750	0.250				
17	45.50	0.700	0.100				
18	46.50	0.650	0.050				
19	47.50	0.600	0.050				
20	51.50	0.550	0.050				
21	56.00	0.500	0.000				
22	58.00	0.400	0.000				
23	59.50	0.250	0.000				
24	61.50	0.100	0.000				

临床流行病学和循证医学

工作簿1 - Excel						
文件 开始 插入 页面布局 公式 数据 审阅 视图 登录						
SUM : $\times$ $\checkmark$ f_x $= (1-B2)^2 + C2^2$						
	A	B	C	D	E	F
1	cut off	se	1-sp	Youden' s index	Euclidean' s index	product index
2	19.00	1.000	1.000	0.00	$= (1-B2)^2 + C2^2$	
3	22.50	1.000	0.950			
4	26.50	1.000	0.900			
5	28.50	1.000	0.850			
6	29.50	1.000	0.800			
7	31.00	1.000	0.750			
8	33.00	1.000	0.700			
9	34.50	1.000	0.600			
10	36.00	1.000	0.500			
11	37.50	0.950	0.400			
12	39.00	0.900	0.350			
13	41.00	0.900	0.300			
14	42.50	0.850	0.300			
15	43.50	0.800	0.250			
16	44.50	0.750	0.250			
17	45.50	0.700	0.100			
18	46.50	0.650	0.050			
19	47.50	0.600	0.050			
20	51.50	0.550	0.050			
21	56.00	0.500	0.000			
22	58.00	0.400	0.000			
23	59.50	0.250	0.000			
24	61.50	0.100	0.000			

临床流行病学和循证医学

工作簿1 - Excel						
文件 开始 插入 页面布局 公式 数据 审阅 视图 登录						
C2 $\times$ $\checkmark$ f_x =B2*(1-C2)						
	A	B	C	D	E	F
1	cut off	se	1-sp	Youden' s index	Euclidean' s index	product index
2	19.00	1.000	1.000	0.00	1.00	=B2*(1-C2)
3	22.50	1.000	0.950			
4	26.50	1.000	0.900			
5	28.50	1.000	0.850			
6	29.50	1.000	0.800			
7	31.00	1.000	0.750			
8	33.00	1.000	0.700			
9	34.50	1.000	0.600			
10	36.00	1.000	0.500			
11	37.50	0.950	0.400			
12	39.00	0.900	0.350			
13	41.00	0.900	0.300			
14	42.50	0.850	0.300			
15	43.50	0.800	0.250			
16	44.50	0.750	0.250			
17	45.50	0.700	0.100			
18	46.50	0.650	0.050			
19	47.50	0.600	0.050			
20	51.50	0.550	0.050			
21	56.00	0.500	0.000			
22	58.00	0.400	0.000			
23	59.50	0.250	0.000			
24	61.50	0.100	0.000			

临床流行病学和循证医学

工作簿1 - Excel

文件 开始 插入 页面布局 公式 数据 审阅 视图 登录

G19

	A	B	C	D	E	F
	cut off	se	1-sp	Youden' s index	Euclidean' s index	product index
1						
2	19.00	1.000	1.000	0.00	1.00	0.00
3	22.50	1.000	0.950	0.05	0.90	0.05
4	26.50	1.000	0.900	0.10	0.81	0.10
5	28.50	1.000	0.850	0.15	0.72	0.15
6	29.50	1.000	0.800	0.20	0.64	0.20
7	31.00	1.000	0.750	0.25	0.56	0.25
8	33.00	1.000	0.700	0.30	0.49	0.30
9	34.50	1.000	0.600	0.40	0.36	0.40
10	36.00	1.000	0.500	0.50	0.25	0.50
11	37.50	0.950	0.400	0.55	0.16	0.57
12	39.00	0.900	0.350	0.55	0.13	0.59
13	41.00	0.900	0.300	0.60	0.10	0.63
14	42.50	0.850	0.300	0.55	0.11	0.60
15	43.50	0.800	0.250	0.55	0.10	0.60
16	44.50	0.750	0.250	0.50	0.13	0.56
17	45.50	0.700	0.100	0.60	0.10	0.63
18	46.50	0.650	0.050	0.60	0.13	0.62
19	47.50	0.600	0.050	0.55	0.16	0.57
20	51.50	0.550	0.050	0.50	0.21	0.52
21	56.00	0.500	0.000	0.50	0.25	0.50
22	58.00	0.400	0.000	0.40	0.36	0.40
23	59.50	0.250	0.000	0.25	0.56	0.25
24	61.50	0.100	0.000	0.10	0.81	0.10

Sheet1

就绪

临床流行病学和循证医学

100%

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](https://www.iikx.com)转发