预测模型好不好，看看SPSS校正曲线就知道

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/1932.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！

临床研究中，我们常采用各种模型预测病人的结局(比如常见的logistic回归模型)。很重要的一点是，要保证预测模型是靠谱的。靠谱的模型才有意义，预测结果才有可能影响临床决策。怎么才算靠谱呢，预测结果和实际观察结果越接近越好。那，怎么知道一个模型的预测结果是有多接近实际观察结果呢?

这里给大家介绍一个洋气的指标，calibration curve或者calibration plot，直译过来就是校准曲线或校准图。其实，校准曲线就是实际发生率和预测发生率的散点图。实质上，校准图曲线是把Hosmer-Lemeshow拟合优度检验的结果可视化了(为方便起见，后面我们简称为H-L检验)。

目前校准曲线常用来评价logistic回归和COX回归模型，以logistic回归为例，通过SPSS在logistic回归界面，点击“选项”，在“统计和图”模块选中“拟合优度”，如图所示。



在logistic回归界面，通过“选项—统计和图——拟合优度”，SPSS就会输出下表。表中步骤1的1到10是根据每个观察单位的预测概率从小到大排序，分成的十组。从表中低出生体重儿的期望值逐渐增加就可以看出来，这里是按照低出生体重儿的预测发生率从小到大排序，并平均分成十组的。表中的实测和期望分别对应的是实际观察结果和预测结果，我们根据这个表即可绘制校准曲线或校准图。

霍斯默-莱梅肖检验的列联表

		分组 = 正常		分组 = 低出生体重		
		实测	期望	实测	期望	总计
步骤 1	1	16	16.697	3	2.303	19
	2	15	15.303	4	3.697	19
	3	18	17.461	5	5.539	23
	4	16	13.215	2	4.785	18
	5	15	13.708	4	5.292	19
	6	15	12.683	3	5.317	18
	7	11	12.310	7	5.690	18
	8	11	12.279	8	6.721	19
	9	6	10.531	13	8.469	19
	10	7	5.812	10	11.188	17

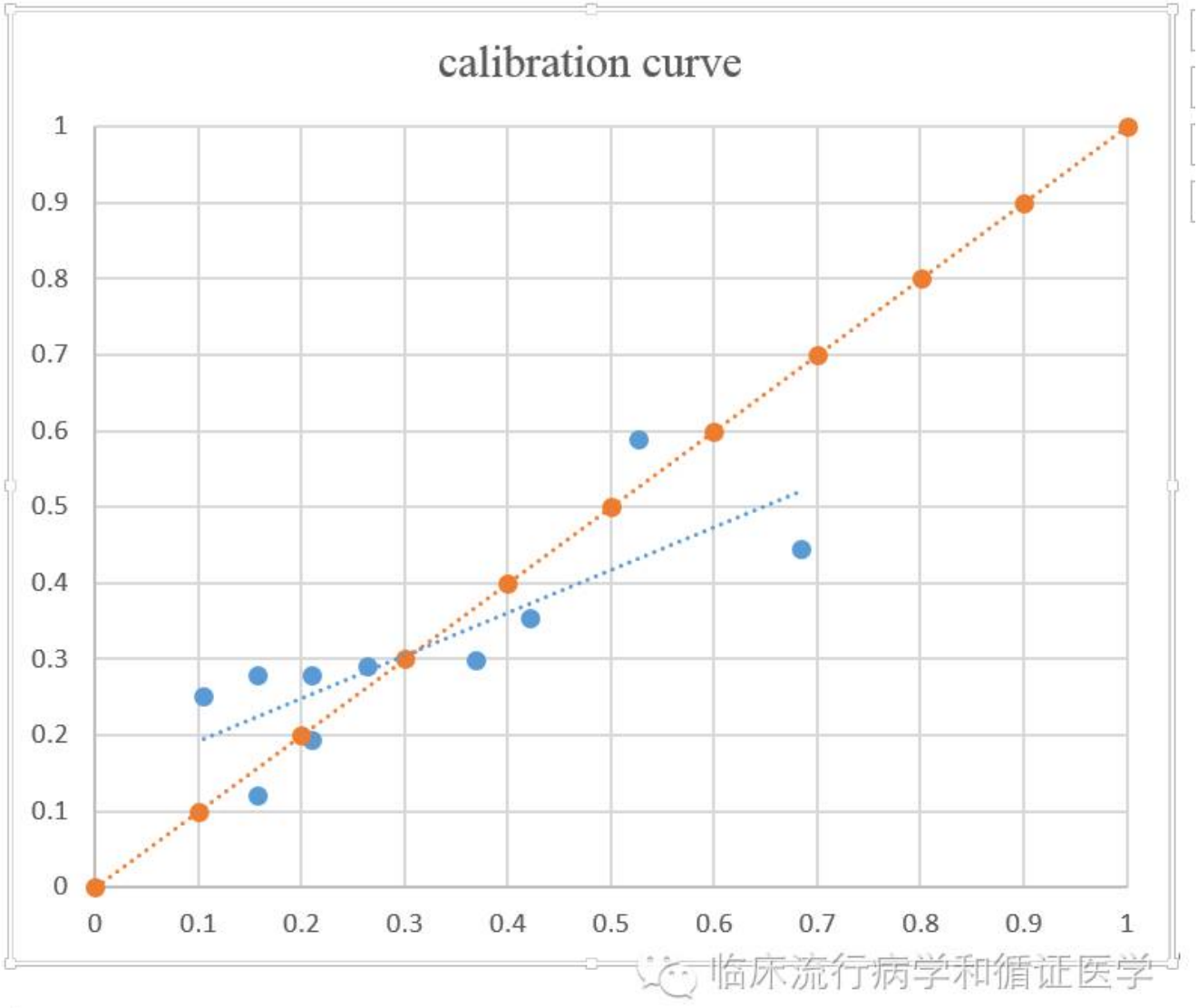
临床流行病学和循证医学

表中给出了因变量的两个取值，正常出生体重儿和低出生体重儿。咱们以低出生体重儿的发生率为例，绘制预测发生率和实际发生率的散点图。根据每组的实测人数、期望人数和每组合计人数，可算出每组的实测发生率和预测发生率，如下图。(小编是把SPSS输出结果贴到Excel里，然后计算了预测和实测发生率，并绘制了二者之间的散点图)

低出生体重儿的发生情况

实测	期望	每组合计	实测发生率=实测/每组合计	预测发生率=期望/每组合计
3	2.303	19	0.157895	0.121211
4	3.697	19	0.210526	0.194579
5	5.539	23	0.263158	0.291526
2	4.785	18	0.105263	0.251842
4	5.292	19	0.210526	0.278526
3	5.317	18	0.157895	0.279842
7	5.69	18	0.368421	0.299474
8	6.721	19	0.421053	0.353737
13	8.469	19	0.684211	0.445737
10	11.188	17	0.526316	0.588842

临床流行病学和循证医学



霍斯默-莱梅肖检验

步骤	卡方	自由度	显著性
1	9.979	8	.266

临床流行病学和循证医学

根据实际发生率和预测发生率绘制校准曲线或校准图。其中蓝线是预测和实测发生率的拟合线，为了方便比较，图中增加了一条理想的橙线，橙线代表的是y=x，代表预测和实测发生率完全一样。所以说，蓝线和橙线越接近说明预测和实测发生率越接近，说明模型越好。

另外，我们说过校准曲线其实是把H-L拟合优度检验可视化了，所以说除了校准曲线，我们还要看H-L拟合优度检验的结果。结果显示，H-L拟合优度检验p值0.266>0.05，说明当前模型和理想中的完美模型没有统计学差异，是可以接受的。也就是说如果校正曲线预测发生率和实测发生率越接近y=x，同时H-L拟合优度检验p值>0.05，说明模型越好。是可以接受的。

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