临床研究中年龄只能是连续变量吗？

作者：writer 来源：临床流行病学和循证医学

本文原地址：https://www.iikx.com/news/statistics/6283.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！

临床研究中年龄只能是连续变量吗？具体的年龄数值可以说把年龄当了连续变量，而80后则把年龄当了分类变量。临床研究中，年龄也是一个很重要的变量，那么问题来了，是把年龄视为连续变量，还是分分段作为分类变量呢?

一个经典的状况(←如果太抽象，请直接看例子)

在探讨危险因素时，我们经常会遇到这样的情况：

a.结局是一个二分类的变量。比如病例对照研究中的“ 病例VS.对照 ”;队列研究中的“ 患病VS.未患病 ”等等。

b.结局存在多个潜在的危险因素。比如性别、年龄、疾病分型、尿蛋白加号的数量等等。

c.希望探讨每个因素是否存在效应以及效应的大小。

这时候您会怎么做呢?

据说百分之好多都这样(←我是错的!我是错的!)

先简单描述一下，另加单因素分析是少不了的。但是最后，肯定还会做一下多因素分析。如果您是老手，您肯定会尝试logistic回归，那么问题来了，是不是把所有变量是直接丢进模型里去就可以了?

随着SPSS等软件的普及，人们对多因素分析模型的使用越来越普遍了，logistic回归模型也成为了应用最广泛的模型之一。我们按照SPSS教程，对着菜单点一点，很容易就能让软件为我们自动拟合出一个模型。我们问过很多医生，大多数医生认为只要变量选对了，丢到模型里去算，然后读结果就好了。实际上这样做会带来很多问题。

一大波“ 栗 子 ”正在接近中

假设A医院有一组接受某手术治疗的患者，大医生LN试图分析他们发生术后并发症的危险因素。(看上去是个队列研究啊喂!当然也可以是病例对照……)患者的结局自然只有两种：术后出现并发症、无术后并发症。我们想找找哪些因素与出现并发症相关，作为一个资深野兽派研究者，我们提出了一个高大上的潜在危险因素——年龄!

因此，这个例子就是，年龄到底是不是发生术后并发症的危险因素呢?

场景一，傻傻分不清楚

在分析年龄这单个因素对术后结局的影响时，您可能会这样做：直接把年龄当做连续变量，比较两种结局患者的年龄均数。

组统计量

	术后并发症	N	均值	标准差	均值的标准误
年龄（年）	无	448	54.96	10.850	.513
	有	352	55.39	9.723	.518

独立样本检验

		方差方程的 Levene 检验		均值方程的 t 检验					
									差分的 95% 置信区间
		F	Sig.	t	df	Sig.(双侧)	均值差异	标准误差差异	下限 上限
年龄（年）	假设方差相等	2.243	.135	-.571	798	.568	-.422	.739	-1.872 1.028
	假设方差不相等			-.579	784.311	.563	-.422	.729	-1.853 1.009

咱一看结果，哎!这个结果好呀，两组均值这么接近(54.96岁 VS. 55.39岁)，p值还这么大(0.568)。然后就顺手得出一个结论，年龄这个因素与术后并发症发生没毛关系。然后代入logistic回归模型(为了栗子清爽，我们暂时只引入年龄一个变量，加入多个变量时与此相似)。结果当然是年龄仍然没啥意义。

方程中的变量

		B	S.E.	Wals	df	Sig.	Exp (B)	EXP(B) 的 95% C.I. 下限	上限
步骤 1[a]	年龄	.004	.007	.327	1	.567	1.004	.990	1.018
	常量	-.458	.386	1.406	1	.236	.632		

场景二，没抖好的小聪明 也许您胆大心细，一心想要探(阳)究(性)真(结)理(论)。然后就想了，年龄这东西，连续变量不一定准啊!对，给丫变成分类的试试。怎么分呢?看看文献都怎么做，咱就怎么分吧! 咔!!!(李楠大吼一声) 尝试变成分类资料探索并没有什么不妥，但是大家经常会参照文献的方法分段，这就不一定总是正确了。当然大多数情况下，考虑临床实际意义进行分类还是必要的，但是在探索阶段，似乎直接这样做就会丢掉好多信息。
在这个例子中，不妨让我们先按10岁一个年龄段划一下试试。分完之后，年龄变成了这个样纸。

年龄段

		频率	百分比	有效百分比	累积百分比
有效	30.00	56	7.0	7.0	7.0
	40.00	164	20.5	20.5	27.5
	50.00	320	40.0	40.0	67.5
	60.00	188	23.5	23.5	91.0
	70.00	72	9.0	9.0	100.0
	合计	800	100.0	100.0	

我们一看!哎，不错啊，年龄变量一下子清爽多了，那咱们再回归一下吧!!!

方程中的变量

		B	S.E.	Wals	df	Sig.	Exp (B)	EXP(B) 的 95% C.I.	
								下限	上限
步骤 1[a]	年龄段	-.004	.007	.352	1	.553	.996	.983	1.009
	常量	-.035	.355	.009	1	.923	.966		

一看结果，估计随之而来的就是失落。得!白折腾半天。神奇的李楠劝大家别着急，咱们马上看看情景三吧。

场景三，低调有内涵 估计您其实是个善于理性思考的医生，在实际分析的时候应该不会出现情景二的后半段。故事的发展应该是接下来这个样子的。在对年龄分段后，您做了一个R×C列联表。

年龄段* 术后并发症 交叉制表

			术后并发症		
			无	有	合计
年龄段	30.00	计数	40	16	56
		年龄段 中的 %	71.4%	28.6%	100.0%
	40.00	计数	80	84	164
		年龄段 中的 %	48.8%	51.2%	100.0%
	50.00	计数	176	144	320
		年龄段 中的 %	55.0%	45.0%	100.0%
	60.00	计数	104	84	188
		年龄段 中的 %	55.3%	44.7%	100.0%
	70.00	计数	48	24	72
		年龄段 中的 %	66.7%	33.3%	100.0%
合计		计数	448	352	800
		年龄段 中的 %	56.0%	44.0%	100.0%

看到表您就发现了，在30岁这个年龄段只有28.6%的人有术后并发症;而40-60这三个年龄段的术后并发症发生率则提高到44.7%-51.2%;到了70岁发生率又回落到了33.3%。从不同年龄段并发症的发生率来看，似乎是一个U字形，并非是完全无关的。也许我们应该将中间三组合并为一个年龄段，保留头尾两个年龄段，这样分组更加合适。先别急，还是让我们先分析分析再说吧。我们先做个卡方检验试试。

卡方检验

	值	df	渐进 Sig. (双侧)
Pearson 卡方	12.369[a]	4	.015
似然比	12.668	4	.013
线性和线性组合	.352	1	.553
有效案例中的 N	800		

a. 0 单元格(0%)的期望计数少于 5。最小期望计数为 24.64。

哎!确实各年龄组间的并发症发生率存在差异!那为什么之前logistic回归却显示年龄与并发症无关呢?其实问题是我们把变量代入模型的方式不对。(你不是玩儿我吧!)没错，就是这个原因。

聊一毛钱儿的 无论年龄是连续变量还是分类变量，直接代入模型的时候，模型都会将年龄与并发症发生风险间的关系想象成等比例递增的关系。换句话说，模型会认为年龄增加1岁，或年龄增加一个年龄段，其发生风险增加的幅度是一致的。也就是说从30岁到31岁与从50到51岁风险增加相同，从30到40岁与从60到70岁风险增加也相同。然而，从上面的交叉表我们就能看出，年龄与并发症发生风险的关系并不是这样的，而是有一个先上升、后下降的趋势。所以在我们分析的时候，就要把年龄变成无序多分类变量代入模型。



然后logistic回归的结果马上就变样了。

方程中的变量

基线，此处是30岁		B	S.E.	Wals	df	Sig.	Exp (B)	EXP(B) 的 95% C.I.	
								下限	上限
步骤 1[a]	年龄段			12.077	4	.017			
40	年龄段(1)	.965	.335	8.323	1	.004	2.625	1.363	5.057
50	年龄段(2)	.716	.316	5.115	1	.024	2.045	1.100	3.803
60	年龄段(3)	.703	.330	4.530	1	.033	2.019	1.057	3.857
70	年龄段(4)	.223	.387	.332	1	.565	1.250	.585	2.670
	常量	-.395	.091	18.701	1	.000	.674		

北医三院临床流行病学循证医学中心

我们看到，年龄这个无序多分类变量还是有意义(p=0.017)的!在各个分组中，40、50、60这三组，与30岁组相比，都是有统计学意义的危险因素，而70岁组与30岁组风险相当。从专业上看，似乎也可以理解，毕竟对于高龄老人，如果术前考虑手术风险较高，医生就会放弃手术了，而真正接受手术的患者，其基本健康状况可能优于一般人群。当然，之后我们还可以把中间三组合并在分析，我们就不详细展示了。

总结 从上面这一斤多栗子中，我们看到的是对于连续变量、有序分类变量进行分析，尤其是代入logistic回归分析时的大致思路。并不是直接比比，然后丢模型就ok了，而是要多探索，结合临床实际对人群进行分组，并结合实际解读分析结果。这样才能够充分的利用数据。 不过需要额外说明的是，所有这些分析上的尝试，仅仅适用于探索性和培育性的研究，也就是说我们希望通过向数据学习，而挖掘一些新的知识。对于验证性的研究，比如药物临床试验，此时所有的研究因素都应该是已知的，研究的目的知识验证假说、得到证据。这是我们应该严格按照统计计划书完成分析内容，而不应该进行上述尝试，否则就有玩弄数据和统计方法、攫取阳性结论的嫌疑了。

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