
论文撰写中常见的统计学问题及其处理

作者：writer 来源：爱科学

本文原地址：<https://www.iikx.com/news/statistics/33.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！

绝大多数的论文撰写,均需通过一定数量临床病例(或资料)的观察,研究事物间的相互关系,以探讨客观存在的新规律。如确定新诊断、新治疗等措施是否优于原沿用的方法,就需进行两种方法比较,这就涉及统计处理;统计设计又是整个课题研究设计中一个重要的组成部分。显然,经正确统计处理的结果可信度高,论文的质量也高。

据不完全统计,在难以发表的、已凝聚着作者心血并花费较长时间与较大财力撰写的研究论文中,约半数以上是由于统计错误致其结果与原文主要结论相违背。如一文采用某新药引产,96例足月孕妇的产后出血与新生儿低Apgar评分率均为2.1%(各2例),明显低于应用原药引产的19例,其产后出血与新生儿低Apgar评分发生率均为15.8%(各3例, $t=7.164, P<0.001$)。故认为采用新药引产是一更安全的措施。原药引产组例数偏少暂且不谈,该资料比较应采用精确法分析,结果是与原结果恰恰相反($P>0.05$),这样上述的主要结论就欠可靠而难以发表,否则论文可起误导作用。类似问题文稿中还常有出现。现就文稿中常见的统计问题及其相应的处理方法简述如下。

一、常用的统计术语

统计学中常用的概念有总体与样本、随机化与概率、计量与计数、等级资料及正态与偏态分布资料、标准差与标准误等。如某研究采用经会阴途径测定宫颈长度,以探讨不同宫颈长度与临产时间的关系。结果显示35例宫颈长度为25~34mm者与32例宫颈长为15~24mm者临产时间的均值 \pm 标准差($\bar{x} \pm s$)各为 57.6 ± 58.1 与 47.3 ± 49.1 小时。该计量资料,经t检验显示 $t=0.780, P>0.05$,并未提示不同宫颈长度的临产时间差异有显著意义;从标准差大于均值,显示各变量值离散程度大,呈偏态分布,故不能采用 $\bar{x} \pm s$ 这一算术均数法计算均数。经偏态转换成近似正态分布资料后结果是:35例与32例的临产时间各为 34.5 ± 4.1 与 26.7 ± 4.1 小时, ($t=7.778, P<0.001$),两组差异有极显著意义。可认为随着宫颈长度的缩短、临产时间也缩短。此外,当两组资料单位不同时,其S单位也不同;即使两组单位相同的变量值,若其均数差异较大,也都应以变异系数替代s来比较两组值的离散度的大小。

二、正常值范围及异常阈值的确定

如何选择研究对象,至少需多少例,正确统计处理和参考一定数量的病例数据,是确定正常值范围及异常阈值的四个重要因素。

1.研究对象:应为“完全健康者”,可包括患有不影响待测指标疾病的患者。如“正常妊娠”的条件:孕前月经周期规则、单胎、妊娠过程顺利、无产科并发症及其它有关合并症,分娩孕周为37~41周+6,新生儿出生体重为2500~4000g和Apgar评分 ≥ 7 分。

2.观察数量:观察数量应尽可能多于100例;需分组者,各组人数也是如此(标本来源困难时酌情减少)。有些指标值如雌三醇(E3)、甲胎蛋白(AFP)、胎盘泌乳素(HPL)等随孕周进展而变化,应按孕周分组;邻近孕周均数相近者,可合并几周计算。若为偏态分布,应以百分位数计算,则例数应 >120例。取各孕周对象时,应考虑到所取各孕周中的例数分布大致均衡。显然,文稿中往往以少量例数求得正常值是欠可靠的。

3.统计处理:应根据所得数据分布特征采用不同的统计处理方法。属正态或近似正态分布的数据,可采用 $\bar{x} \pm s$ 法计算;这也适用于以一定方法能将非正态分布转换成正态或近似正态分布的资料。对无法转换的偏态资料,应采用百分位数计算法。具体计算(包括上下限初步制定)见文献。

4.对照数量:相应观察的病例数(包括分组)应不少于30例,这对制定某指标有临床意义的异常阈值尤其重要,这一点往往易被忽视。如在参考较多病例数据后,唾液游离E3的下限异常阈值应为第2.5百分位数,而非通常采用的5百分位数。否则,将会导致该指标产前监护的假阳性率增加。

三、t检验与校正t检验(t_c检验)

这是文稿中极易混淆的一类计量资料统计问题。

(一)检验的注意事项

1.t检验的意义:t检验与所有统计分析相同,其结果提示现有差别不仅仅是抽样误差所致,且提示犯第一类错误的可能性大小,即 $t_{0.05}$ 与 $t_{0.01}$ 犯第一类错误的可能性各为5%与1%。

2.统计意义与临床意义的关系:统计学有显著意义,而在临床上可能是无意义的,提示该研究应继续深入,以明确该差异是否真有显著意义;相反,统计无显著意义,而临床上却是有意义的,不能贸然轻易地下结论。应复查实验设计、方法、试剂及仪器性能、质控措施和实验数据等是否有问题,或尚需再进一步增加样本量进行复测等。

3.t检验适用范围:t检验仅适用于正态或近似正态分布(包括偏态转换)和其方差是齐性资料的检验;t_c检验适用于可比性资料,即除了欲比较的因素外,其它所有可影响的因素应相似。

4.t检验的结果判断:判断结果不应绝对化, $P < 0.05$ 或 $P > 0.05$,分别表示可拒绝或接受原定的假设,但两者都有5%的可能性犯第一类错误;而P值越小,只能是更有理由拒绝原定的假设。

5.单侧与双侧检验:应预先制定本研究的结果是需行双侧还是单侧检验。对有把握确知某治疗措施或某指标是不会劣于现有的,才作单侧检验;若不知何者为优,应行双侧检验。因为在同一t值的界限上,单侧检验的概率(P)仅为后者的一半,也就是说单侧检验较双侧检验更易得出差别有统计意义的结论,不可随意制定。一般讲,绝大多数研究以采用双侧检验为妥。

(二)t_c检验与t检验的区别

当两样本均数的方差非齐性时,应以t_c替代t检验。例如:甲组32例血清某指标值为 53.9 ± 49.6 ($\mu\text{mol/L}$);乙组6例的结果为 26.6 ± 7.2 ($\mu\text{mol/L}$),若不考虑两样本方差大小,t检验示 $t = 1.331, P > 0.05$,提示两组血清该指标的平均含量差异无显著意义。但先作方差齐性检验, $F = 47.4, P < 0.01$,示这两样本方差差异有极显著意义。据此应采用t_c检验, $t_c = 2.952 > t_{0.012, 875}, P < 0.01$ 。显然,与上述结论恰恰相反。

四、卡方(χ^2)、校正 χ^2 与直接概率法(或精确法)检验

这三种检验方法为一类用途较广、但也易混淆的、适用于计数资料检验的方法。应注意,鉴于总数与理论值的不同,应采用相适合的检验方法。

例1.192例出生体重 $< 4000\text{g}$ 的新生儿发生难产与窒息数分别为151例与22例;3475例出生体重 $3500 \sim 4000\text{g}$ 的新生儿发生难产与窒息数分别为185与265例;2451例出生体重 $2500 \sim 3500\text{g}$ 的新生儿发生难产与窒息数分别为122与169例。3组的构成比:难产与新生儿窒息率分别为:78.6%、5.3%、5.0%与11.4%、7.6%、6.9%。据此贸然认为出生体重 $2500 \sim 3500\text{g}$ 为最佳新生儿分娩体重的结论是不可靠的。经 χ^2 分析,后两组的难产与窒息率间和前两组窒息率间差异均无显著意义(P 均 >0.05)。故可认为,单据本研究结果是难以得出上述临床上认可的结论的。这涉及到上述“统计无显著意义,而临床却是有意义”的问题,应进一步复查或增加样本测试。杜绝单纯根据百分率的大小贸然下结论。

例2.某药治疗感染衣原体(CT)的中、晚期孕妇各11例和36例,她们的新生儿感染CT数各为3例和23例。 χ^2 检验得 $\chi^2=4.570, P<0.05$ 。据此误认为,某药治疗中孕期感染CT孕妇的新生儿感染CT数少于晚孕期才开始治疗的新生儿感染数。根据统计原则,其中一个数的理论值为4.9(<5)时,应采用校正 χ^2 计算,得 $\chi^2=3.209, P>0.05$ 。显然,正确结论恰与上述相异。

例3.以精确法替代 χ^2 检验。某新技术测试8例卵巢内胚窦瘤患者,5例呈阳性反应;测试25例卵巢颗粒细胞瘤患者中6例阳性。 χ^2 检验得 $\chi^2=4.042, P<0.05$ 。误认为该新技术测前组的阳性率高于后组。但鉴于总例数33例(<40),且其中一个数的理论值为2.7(<5),故应改用精确法检验,结果首次计算 P 值,已达0.102, $>$ 双侧检验的有显著性意义的界限0.025,故 $P>0.05$ 。结论也恰相反。

五、相关与回归分析

相关分析只是以相关系数(r)来表示两个变量间直线关系的密切程度和相关方面的统计指标。无论是正相关(r 为正值)或负相关(r 为负值),只是经相关系数的统计意义检验(如 t 检验)后,当 $P<0.05$ 时,即示差异有显著意义时,才能依据 $|r|$ 值的大小来说明两变量间相关的密切程度。因此,表示相关性,除写出 r 值外,应注明 P 值;切不可将相关的显著性误解为相关程度;也应注意:相关分析是不能单纯用于阐明两事物或现象间存在着本质的联系,即使两变量间存在高度相关关系(即有一定的统计联系),也不能证明它们间存在着因果关系。如欲证明两事物间的内在联系,必需凭借专业知识从理论上加以阐明。

“相关”是表示两个变量间相互关系的密切程度,而回归分析是提示两个变量间的从属关系。在回归分析中,应注意由 X 变量值推算 Y ,与以 Y 变量值推算 X 的回归线是不一样的;直线回归方程的适用范围,一般仅适合于自变量 X 原测数据的范围,故绘制回归线时, X 值切不可超越实测值的范围而任意延长。

可见,这两种分析,说明的问题是不同,但相互又有联系。在作回归分析时,一般先作相关分析,只有在相关分析有统计意义(即回归有统计意义)的前提下,求回归方程和回归线才有实际意义。决不能把毫无实际意义的两个事物或两种现象进行相关与回归分析。

六、数据的正确书写

1.文稿内各数据的书写必须前后一致;总数应等于各分组的数据之和。

2.对不同指标,有其不同数据精度的要求,这应结合专业知识加以判断。如新生儿出生体重是以公斤为单位,记录测定数据精确到小数点后的第二位数字即可。

3.测定数据的书写,不能超越其测量仪器测试的精确度范围。

4.同一指标的前后数据应保持同一精确度。

5.经计算,出现比预定小数点后两位数多的数字,应采取“ 4舍、 6入”与“5‘奇’进‘偶’出”方法,以决定小数点后第三位数字是“舍”还是“入”,即5前为单数则入,双数则舍。

6.未经统计检验,文稿内不宜出现推断性的比较结果的结论。如“××结果的百分率高或低于××结果的百分率”、“本文结果较××报道的多或少或类似”等结论。这在综述类文章撰写过程中也需注意;欲予以比较,也应注意两者的可比性。

上述是统计学中较为基础的概念,但又是文稿中常见的、较易出现差错的内容。因此可以认为,统计学是医护人员必需掌握并能熟练应用的一门重要知识;藉此,可不断地从自身和他人的研究获取更多、更新和更可靠的专业信息。

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://www.iikx.com)转发