

有关 LMS 软件程序应用介绍

蒋一方, 林钟芳 (上海交通大学附属儿童医院, 上海 200040)

中图分类号: R174 文献标识码: B

摘要: 介绍构建儿童体质指数 (body mass index, BMI) 的方法—LMS 法软件的使用过程, 以促进我国儿童体质指数及其它参数的百分位数的研究工作。

关键词: LMS; 软件; 体质指数

2000 年国际肥胖工作小组 (International Obesity Task Force, IOTF) 推出了采用英国伦敦大学儿童健康研究所 Cole 教授创立的 LMS 方法构建儿童体质指数 (body mass index, BMI) 百分位曲线^[1]。在之前他本人已用此法建立了英国青少年儿童的身高、体重、头围、腰围上臂围及 BMI 百分位曲线^[2]。我国有些城市如上海、西安等地也采用了此种方法建立了 0~18 岁青少年儿童身高、体重及 BMI 百分位曲线^[3,4]。因此让更多研究人员了解 LMS 方法及其应用, 将对我国 BMI 及其它参数的百分位数研究有重要的促进作用。目前已有 LMS 方法的应用软件, 本文详细介绍 LMS 方法软件的应用。

所谓 LMS 程序, 就是应用 LMS 方法对有关生长发育测量值、血尿化验数据或其它生理指标进行处理, 使之成为光滑的百分位曲线。百分位曲线显示出一项测量值的群体分布情况, 并且会随某些变量而相应变化, 如年龄因素。LMS 方法综合了三条曲线而改变其分布, 这三条曲线分别是中位数曲线 (M)、变异系数曲线 (S) 和不对称曲线 (L)。L 曲线使测量值由偏态资料经 Box-Cox 转化为正态的幂 (λ) 随协变量变化 (如年龄) 的一条曲线。使用惩罚似然法使这三条曲线达到光滑。可以选择合适的光滑参数或有效自由度 (effective degree of freedom, edf) 来获得。LMS 软件的操作程序主要包括五个步骤:

1 资料输入 (data entry)

资料的输入及输出通过文件菜单来操作。先将资料输入 Excel, 以文件 A 保存。然后建立新的 Excel 工作表文件 B, 将文件 A 中的原始数据复制后, 以选择性粘贴“数值”的方式, 粘贴到新的工作表, 其中首列为年龄, 第二、第三列为测定值 (身高、体重), 第四列为 BMI。这种安排是 LMS 程序所能接受的, 也是理想的。然后将 Excel 工作表 B 以 *.txt 文件保存。

2 模型调整 (model fitting)

点击 Model, 在对话框内进行 Age Options 选择, 共有三种: 原始年龄 (original age, O)、重设年龄 (rescaled age, R) 或转换年龄 (transformed age, T) 以及 edf 值的选择。在建立体重和身高模型时, 可选择重设年龄。在建立 BMI 模型时, 可选择转换年龄。当选择合适时, 将自动出现四条曲线: L 曲线、M 曲线、S 曲线和百分位曲线。同时 LMS 程序还能显示 Detrend QQ 图来测试模型调整的正常化, 就匹配的好坏可以从点击 Graph, 在菜单中选择 QQ Tests 来了解 L、M、S 三条曲线的数值, 一般应在 2 以下较好。点击 Graph, 还可选择 Z scores, Data plot (双击该图会出现百分位曲线)。模型调整可能要做许多次方能达到要求。LMS 程序可以提供目前建立的模型和前一

模型百分位曲线的比较。

2.1 何谓 edf edf 代表每一个 L、M 和 S 曲线的复杂程度。Edf 为 2 相应于直线, edf 为 3 可给出一简单曲线, 如二次曲线, 以及 edf 为 4 或更大提示曲线更复杂。

2.2 选择 edf 值的命令 ① 通过增加或减少 edf 值的一个单位, 最佳化 M 曲线在变异方面的改变是小的, 小样本 ($n < 500$) 将 edf 设置为 5 是合适的。大样本 ($n > 10\,000$) edf 值可能达到 15 或 15 以上。④ 在大部分情况下, 最佳化 S 曲线将 edf 设定为 3 将是足够的, 虽然 edf 可高达 10 或 10 以上, 对大样本可能是需要的。④ 最佳化 L 曲线, 设定 edf 为 3 在许多情况下为太大, 可假设为 0, 此时无斜率改变。然后设定为 1, 在所有年龄是一条直线的调整, 然后每增加 1 作一次观测。超过 6 是不可能的, 除非对大样本资料, 其覆盖年龄范围很广。

2.3 选择 edf 值一般规则 设定 L、M 和 S 的 edf 值分别为 3、5 和 3, 可按下列步骤调整: ① 每次改变 edf 值一个单位; ④ 了解曲线改变的形状; ④ 为了解曲线形状随 edf 值变化而变化, 可双击曲线图, 窗口可出现先前 (蓝色) 和目前 (黑色) 图像比较; ④ 在模型窗口观察变异的改变, 在 edf 中小于 4 个单位的改变是小的改变。改变在 10 或 10 以上是大的改变; ④ 如果不能确定 edf 值时, 应从小的开始; ④ 大样本需要大的 edf 值; ④ 一般来说, $M\ edf > S\ edf > L\ edf$; ④ 如果你有 LMS 程序, 可以使用 Detrend QQ 图和 QQ Tests 来调整。

2.4 年龄选择 ① 原始年龄 (O): 是指原始数据上的年龄即实际年龄。④ 重设年龄 (R): 根据调整 M 曲线形状而作的一种经验转换。M 曲线首先用原有年龄来调整, 然后绘制出 M 曲线的形状, M 曲线的斜率与转换年龄是一条直线, 当该 M 曲线根据转换年龄调整后, 原有年龄就被储存起来, 这个过程可以大大改进复杂的非单一性资料的调整。请注意如果资料是非单一的, 这个过程可能会导致被调整曲线在不应该弯曲的点出现假象的波动。④ 转换年龄 (T): 幂的转化是 $age + offset$ (分支), 幂 (λ) 的值处于 +3 和 -3 之间, 当幂接近于 0 时, 它是对数转换; 当幂是 0 时, 或者任何值的一个倒数时, 此时 $age + offset$ 一定是正数。点击 transformed age, 可产生一个 power 转化对话框, 并在空格里为 offset 选择一个数据。

3 图像显示 (graphical display)

点击 Graph, 可以选择各种曲线图。LMS 简易版本有六幅图: Data plot, Z scores, L、M、S 和 Centile, 当模型调整成功时, 可显示出 L、M、S 和选择好的 Centile。LMS 程序由两个额外图像也可用作模型校对: Detrend QQ 图和 QQ Tests。当资料

太少时, Q Q Tests 可能不会显示。百分位曲线可以添加到 Data plot, 也可以添加到 Z score plot, 方法是双击这些图像。

有时发现有些数据需要核实(太高或太低), 可先活化 Data plot 窗口, 选中 Identify 后, 点击“ON”按钮, 再按“OK”, 此时点中 Data plot 图中的数据点, 该数据点会变为红色, 在活化窗口左侧的 Data 窗口中显示该点数据(蓝色背景), 对该数据可以进行重新编辑或删除, 点击 back space 键, 该项数据会消失, 然后将 ON 状态转为 OFF 状态后, 图像也随之调整。如果需要重新写图示以便获得合适的较清晰的标记或名称, 可以点击 Edit/ Graph, 选择 Title 选项卡, 在其中可以插入和编辑图像名称、纵轴和横轴名称。同时点击 Edit/ Graph, 还可选择 Scale 选项卡, 点中 User defined 后, 可建立每根轴的最大值、最小值和刻度间距。注意如果处于 On 的状态时, 不允许双击 Data plot, 或 Z score plot, 来除去或添加百分位曲线。也就是说, 当点核实后, 应将 On 状态转化为 Off 状态。有时候, 所选择的 Age Options 以及 edf 值并不能建立合适模型, LMS 程序会提出警告, 可能需要核对某些点的数据, 删去后可能建立模型。

4 模型核实(model checking)

模型窗口指示出目前模型的调整结果, 是由偏差(deviance)来测量的, 这个概念本身难以解释。当模型发生改变时, 模型窗口显示出 edf 值的改变以及相应 deviance 的变化, deviance 的变化是接近于 edf 值改变的卡方分布, 这提供了一种简易的方法来判断模型改变的重要型。从理论上讲, edf 为 1 时, 可使 deviance 改变 4 个单位, 这在 5% 水平上正好是有意义的。实际操作中, deviance 的变化需要合理的增大, 在它变得重要之前可能是 8 或者更大。对于非常大的样本($n > 10\ 000$), deviance(> 20)大的变化可能使所调整曲线的形状只有较小的改变。所调整模型的一种累积对数, 它们的 deviance 和参数可以在 monitor dialogue 中发现。下面的核对将是有用的: ① Z score 图将对核对结果有用; ②目前和先前的比较, 如 L、M、S 和 Centile 曲线的比较将是有用的; ③ Detrend Q Q 图是 Q Q 图的一种 Detrend 形式; ④ Q Test 有助于试验模型调整的好坏, 并选择最后的模型。一个最后的模型可以这样来考虑: 如果对于 L、M 和 S 而言, Q 统计曲线处于 10% ~ 90% 范围以内。

一旦一个模型已经调整妥当, 该资料被转化为 Z score, 可以点击 Graph 选择 Z score 图。在 Z score 图上双击, 可显示隐藏的百分位曲线图。该百分位曲线是否是一条水平直线, 同时点击 Z score 图使之活化, 然后可以进行太大或太小值的点的核实, 其方法和 Data plot 一样, 点击 Identify 来核实被怀疑的点。一般来说, 所采集的资料要进行预处理, 它包括两个步骤。第一, 是删去明显不在总体分布的散在点, 此过程称之为删去异常值。第二是从 Z score 图中删去大于或等于 + 5s, 或小于或等于 - 5s 的点, 这项工作可能要重复进行 2 ~ 3 次。注意, 显示百分位曲线和核实 Data points 不要同时操作, 如果 Identify 处于“on”状态, 双击该图并不能显示百分位曲线。

5 模型保存(model saving)

可以通过菜单来保存一个模型。通常可以采用两种方法: ① Save as 方式保存, 该方法是保存整个模型和原始数据(或改动过的数据), 文件形式为 *. LMS 或 *. lms; ② Export 方式保存, 该方法是保存 L、M、S 和百分位数值(没有图), 文件形式为 *. EPT。

LMS 系统文件: 一个系统文件可由 LMS 或 lms 的扩展名组成。例如, test030503t. lms 或 TEST 030503T. LMS。LMS 程序提供一个首选的名称, 当保存文件时, LMS 系统文件含有三个部分: (1) 最初部分包括在调整模型二条直线中的一些参数。最初一条线提供 LMS 曲线的 edf 值, 第二条线包括选择年龄的值, 所用资料的大小以及加权和选择的状态。LMS 程序将增加在直线的最后加上组的数目。(2) 第二部分提供年龄及相应的所调整模型的 L、M、S 曲线。(3) 第三部分包括年龄、测量值、Z scores(LMS 简易版本), 或者在 LMS 程序中已选择加权与所选的列。所有的资料本包括在内, 并且被删去的数字用“*”号标记出来。

注意: 建议不要编辑一个系统文件, 因为它含有的信息将可用予以 LMS 重新恢复以模型的特性。如果你确实需要编辑一个文件, 可将文件复制后以 txt 文件(文本文件, 以制表符分割)来编辑。

在运行 LMS 程序开始时, 就要设定好百分位曲线的位数, 如 9 项: P₃、P₅、P₁₀、P₂₀、P₃₀、P₈₀、P₉₀、P₉₅、P₉₇, 也可以选 7 项。其方法是先点击 Edit, 选择 Preference, 然后点击 Centile, 或不等的标准差距(0.666 7)。如选择 Centile, 然后便显示 Select centile, 根据需要可填写 7 项或 9 项。

致谢: 衷心感谢 Tim Cole 教授和潘蕙琦博士在 LMS 软件应用方面的热情指导。

[参 考 文 献]

- [1] COLE TJ, GREEN PJ. Smoothing reference centile curves: the LMS method and penalized likelihood[J]. Stat Med, 1992, 11: 1305-1319.
- [2] COLE TJ, BELLIZZI MC, FLEGAL KM, et al. Establishing a standard definition for child overweight and obesity worldwide: international survey[J]. BMJ, 2000, 320: 1240-1243.
- [3] 蒋一方, TIM COLE, 潘蕙琦, 等. 上海市 0~18 岁体质指数百分位曲线及超重肥胖界值点标准的研制[J]. 中国儿童保健杂志, 2004, 12(6): 461-464.
- [4] 尚磊, 徐勇勇, 江逊, 等. 陕西省青少年儿童 BMI 指数与美国、荷兰标准的比较[J]. 中国儿童保健杂志, 2002, 10(6): 361-363.

作者简介: 蒋一方(1944), 男, 上海人, 研究员, 硕士学位, 主要从事儿童营养研究。

收稿日期: 2005-02-19

《中国儿童保健杂志》2005 年继续医学教育试题(4)

【信 息】

- 1 媒介对儿童发育的影响有哪些?
- 2 卫生部门在儿童早期综合发展中的作用?
- 3 青春期心理发育的特点?
- 4 按照 WHO 定义, 基本的生活技能包括哪些?