

# 汇报讲稿

2020年12月14日 12:54

- 接下来由我来介绍检验方法的随机模拟与案例研究等内容；
- 首先，进行了相应的一些模拟设置；
  - .....《slide上》
  - 所以模拟的是**这三种情况下**（均值为一个常数，服从均匀分布的协变量与均值相关，服从正态分布的协变量与均值相关），不同样本量情况下，不同检验方法之间的差异。
- 这里的代码较多，所以就没放在slide上面；
- 首先是均值为常量：
  - 构建了一个理论值与第一类经验误差的关系图（P-P plot），用于性能的综合评估；
  - 对于性能良好的测试，P值应接近其相应的经验I型误差，因此曲线应接近对角线。曲线偏离对角线表示性能差，偏离越大，性能越差。由于标称I型误差通常很小，我们将重点放在图的(0, 0)端来评估性能
- 常数均值：
  - Wald的表现最差，因为Wald检验的曲线位于(0, 0)端的对角线之下，并且严重偏离对角线。wald更有可能拒绝 $H_0$ ，因此会产生膨胀的I型错误。
  - LR检验在所有测试中表现最好，因为图接近对角线。
  - 当样本量较小且检测不足的比例较大时，he test 和 score test的表现不如LR测试。he test 和 score test的图位于对角线下方，这表明这两个检验更有可能拒绝 $H_0$ 并产生膨胀的I型错误。
  - 随着检测下的比例变小(割从1.5增加到2.0)，两个测试的性能变得更好。
  - 当平均值中没有协变量时，新测试和分数测试的图是相同的，这由定理2证实。
  - 正如预期的那样，测试的整体性能取决于样本大小和检测不足的比例，样本大小越大，检测不足的比例越小，性能越好。
- $U(0, 1)$ 
  - 同样，wald表现最差，第一类错误急剧膨胀，尤其是当检测比例较大时。
  - LR test表现最好，图接近对角线。
  - he test 和 score test的表现相似。
  - 总的来说，he test 和 score test的图略低于对角线，这表明第一类错误略

有夸大。

- $N(0, 2)$ 
  - 在这种情况下，he test 在四个测试中表现最好，因为它的图接近对角线。
  - 其他测试的图严重偏离对角线。
  - LR test的图位于对角线下方，表示膨胀的第一类错误，而score 和 wald的图位于对角线上方，表示收缩的第一类错误
- 在mTobit模型下，零假设 $h_0$ 不成立，即存在一个被检测的潜在类。检查评估检测潜在类别的能力的power（统计功效），从而正确地拒绝具有更高能力的 $H_0$ 以获得更好的性能。
- 常数均值
  - 图总结了以5%的第一类错误拒绝零假设的经验power。
  - 总体来说，LR测试产生的功率最小，而Wald测试产生的功效最大。
  - 然而，由于Wald产生膨胀的第一类错误，沃尔德检验更有可能拒绝零假设，更大的power并不反映沃尔德检验的良好性能。
  - 由于不涉及协变量，新的和分数测试产生相同的power。
  - 总的来说，当 $\mu$ 是固定的时，功效随着潜在类别的比例和样本量的增加而增加；
  - 功效受Tobit分量中检测数据的比例影响。检测到的数据越少，测试在检测潜在类别方面就越强大。
- $U(0, 1)$ 
  - 一个是 $\omega$ 是一个概率常数，另一个是 $\omega$ 取决于x协变量；
  - 首先是 常数 $\omega$ ：
    - 由于沃尔德测试的第一类错误不可靠，我们将重点放在其他三个测试上。
    - 在所有情况下，LR测试都不如he test的和score test强大。
    - 与score test相比，he test在某些情况下获得了更大的功率。总的来说，he test至少和score test一样强大。
  - $\Omega$ 取决于x协变量：
    - LR测试产生的功率最小；然而，he test比score test表现更好。在某些情况下，he test比score test具有更大的power。
- $N(0,2)$ 
  - $\omega$ 常数：
    - 分数和LR测试产生相似的力量，但是新的测试没有分数和LR测试表现

好，它是最没有力量的。

- Omega取决于x协变量：

- 新测试实现了最大的功效，并且比LR和分数测试具有显著更大的功效。
- 分数测试产生的功效最小，但是LR和分数测试之间的差异没有新测试大。

- 案例研究：

- 在上述模拟研究中，只考虑了一个协变量。
- 为了评估其他可能相关协变量的测试性能，使用多个连续变量和二元变量进行了更多协变量的研究；
- 数据介绍： .....