

基于统计学习方法的 NIPT 时点选择与胎儿的异常判定决策模型

摘要

针对问题一，

针对问题二，

针对问题三，

针对问题四，

关键词：

一、问题重述

1.1 问题背景

NIPT 是一种通过采集母体血液、检测胎儿的游离 DNA 片段、分析胎儿染色体是否存在异常的产前检测技术，目的是通过早期检测确定胎儿的健康状况。根据临床经验，畸型胎儿主要有唐氏综合征、爱德华氏综合征和帕陶氏综合征，这三种体征分别由胎儿 21 号、18 号和 13 号“染色体游离 DNA 片段的比例”是否异常决定。NIPT 的准确性主要由胎儿性染色体浓度判断。通常孕妇的孕期在 10 周 25 周之间可以检测胎儿性染色体浓度，且如果男胎的 Y 染色体浓度达到或高于 4%、女胎的 X 染色体浓度没有异常，则可认为 NIPT 的结果是基本准确的，否则难以保证结果准确性要求。同时，实际中应尽早发现不健康的胎儿，否则会带来治疗窗口期缩短的风险，早期发现风险较低；中期发现风险高；晚期发现风险极高。实践表明，男胎 Y 染色体浓度与孕妇孕周数及其身体质量指数（BMI）紧密相关。通常根据孕妇的 BMI 值进行分组分别确定 NIPT 的时点。由于每个孕妇的年龄、BMI、孕情等存在个体差异，对所有孕妇采用简单的经验分组和统一的检测时点进行 NIPT，会对其准确性产生较大影响。因此，依据 BMI 对孕妇进行合理分组，确定各不同群组的最佳 NIPT 时点，可以减少某些孕妇因胎儿不健康而缩短治疗窗口期所带来的潜在风险。为了研究各类孕妇群体合适的 NIPT 时点，并对检测的准确性进行分析，附件给出了某地区孕妇的 NIPT 数据。在实际检测中，经常会出现测序失败的情况。同时为了增加检测结果的可靠性，对某些孕妇有多次采血多次检测或一次采血多次检测的情况。试利用附件提供的数据建立数学模型研究如下问题：

1.2 问题提出

问题 1：分析胎儿 Y 染色体浓度与孕妇的孕周数和 BMI 等指标的相关特性，给出相应的关系模型，并检验其显著性。

问题 2：男胎孕妇的 BMI 是影响胎儿 Y 染色体浓度的最早达标时间的主要因素。要求对男胎孕妇的 BMI 进行合理分组，给出每组的 BMI 区间和最佳 NIPT 时点，使得孕妇可能的潜在风险最小，并分析检测误差对结果的影响。

问题 3：综合考虑身高、体重、年龄等多种因素对男胎 Y 染色体浓度达标时间的影响、检测误差和胎儿的 Y 染色体浓度达标比例，根据男胎孕妇的 BMI，给出合理分组以及每组使得孕妇潜在风险最小的最佳 NIPT 时点，并分析检测误差对结果的影响。

问题 4：以女胎孕妇的 21 号、18 号和 13 号染色体非整倍体（AB 列）为判定女胎是否异常结果，综合考虑 X 染色体及上述染色体的 Z 值、GC 含量、读段数及相关比例、BMI 等因素，给出女胎异常的判定方法。

二、 问题分析

2.1 问题一分析

2.2 问题二分析

2.3 问题三分析

2.4 问题四分析

三、 模型假设

四、 符号说明

五、 模型建立与求解

5.1 问题一模型的建立与求解

5.1.1 数据预处理

判断空数据，异常数据除去

5.2 相关性分析

首先认为存在相关性，叙述相关性，代入数据算出相关性，根据结果描述相关性强弱。

相关系数皮尔逊斯庇尔曼肯德尔相关系数效果皮尔逊最好
低于 0.05 显著性比较强

5.3 关系模型

5.4 显著性检验

p 检验 t 检验

散点图热力图

由图1可知，Y 染色体浓度与孕周、BMI 均存在一定的相关性，且孕周与 BMI 之间也存在一定的相关性。

1. 数据点集中在孕周 10-20 周，Y 染色体浓度多在 0.05-0.15 之间，显示出随孕周增加的趋势，但分布较集中。
2. BMI 在 25-40 时浓度分布较广，>40 时浓度偏低，提示 BMI 可能影响浓度达标时间。
3. 数据点主要集中在 BMI 25-40 和孕周 10-25 之间，分布较密集。

4. 相关性较低（接近 0），但孕周 10-20 和 BMI 25-35 区域数据密集。

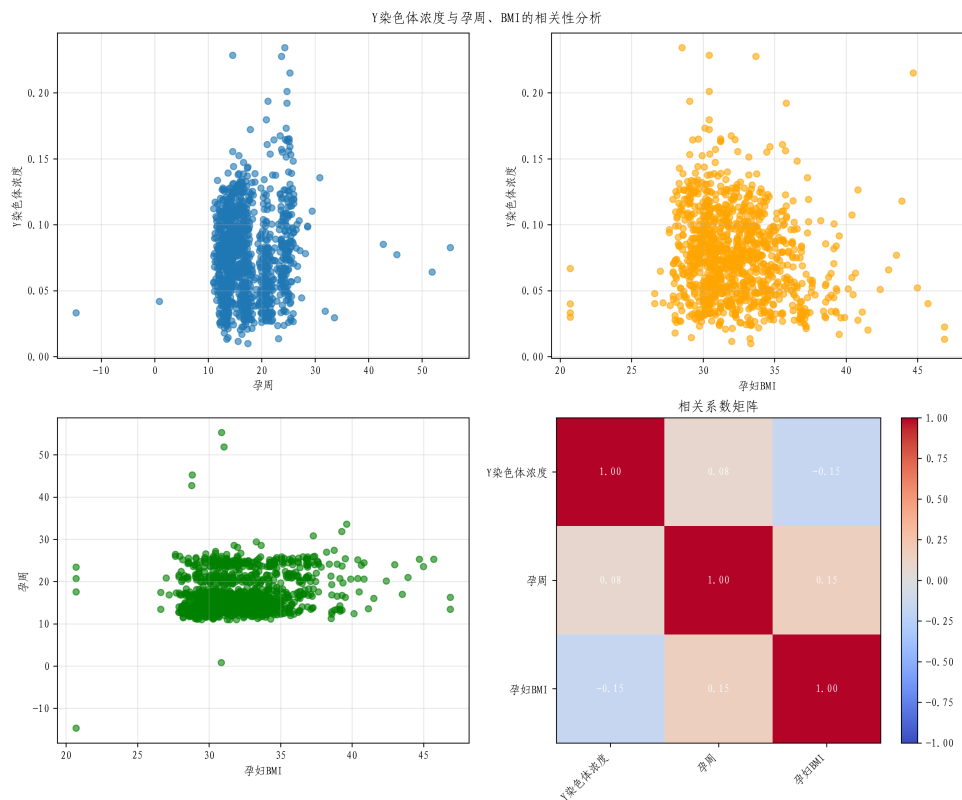


图1 Y染色体密度与孕周、BMI散点图

5.5 问题二模型的建立与求解

5.6 问题三模型的建立与求解

5.7 问题四模型的建立与求解

六、 模型评价

6.1 模型优点

6.2 模型缺点

参考文献

[1] 卓金武. MATLAB 在数学建模中的应用[M]. 北京: 北京航空航天大学出版社, 2011.

[2] 司守奎, 孙玺菁. 数学建模算法与应用[M]. 2 版. 北京: 国防工业出版社, 2015.

[3] 同济大学数学系. 高等数学[M]. 8 版. 北京: 高等教育出版社, 2014.

- [4] REITZ K, SCHLUSSER T. Python 编程之美：最佳实践指南[M]. 电子工业出版社, 2018.
- [5] MITCHELL T. 机器学习[M]. 机械工业出版社, 2008.
- [6] RASHID T, 林赐. Python 神经网络编程 Make Your Own Neural Network[M]. 人民邮电出版社, 2018.