

基于马尔科夫决策的生产过程决策问题

摘 要

本文针对某企业在生产电子产品过程中所面临的次品率、检测及拆解等一系列决策问题展开研究。通过优化生产过程中的抽样、检测和拆解策略,最大化企业利润和生产效率。本文的研究旨在为企业提供一套基于数学建模的决策方案,帮助其在面对不同生产条件时合理选择抽样、检测和拆解策略,提升资源利用率和经济效益。针对问题一,为通过抽样检测判断次品率是否超过标称值,本文分别针对 (1)(2) 情形建立了假设检验问题,进而得到了是否接收零配件的条件。为确定合适的抽样样本量,本文利用检测花销与假设检验效果 (运用假设检验功效函数衡量) 之间相互制约的关系,建立决策函数,得到了最优样本量。针对问题二,为针对六种不同情况做出决策,本研究运用二项分布和几何分布的性质,在不同决策方案下,用各零件检测成本、次品率、调换成本等值表示出每生产一个合格成品的利润期望值。再依次将六种情况的具体数值带入到表达式之中,将利润期望最大化作为决策标准,得到六种情况下的最优决策。针对问题三,为解决多工序制造系统中的零件、半成品与成品的次品检测和拆解决策问题,求解目标在于合理设计检测与回收策略及最大化系统的期望利润。首先构建状态-决策模型,定义零件、半成品和成品的状态变量与决策变量,将问题转化为多阶段决策问题。模型通过动态规划算法有效求解最优决策路径,计算各个阶段的成本和收益,并利用记忆化存储提高求解效率。最终求解结果显示,按照最优检测与拆解策略,系统的最优期望利润为 56.7 元。针对问题四,为描述次品率的不确定性,本文运用贝叶斯估计,选用共轭先验分布 beta 分布,推导出次品率服从的后验分布。完成问题二,计算出在次品率服从的分布下利润的期望,重新进行决策。完成问题三,修改固定的次品率为分布密度后,重新运行程序获得最优决策路径和相应最优利润期望为 54.2 元。经过灵敏度检验,先验分布超参数的选取对利润期望的影响较小。

关 键 词: 假设检验 状态-决策模型 动态规划算法 贝叶斯估计 共轭先验分布

一、 问题重述

1.1 问题背景

为应对人口老龄化及出生人口下降等问题，我国出台了三孩政策。在相关政策实施后，女性的整体妊娠年龄呈上升趋势，而随着妊娠年龄的提高，胎儿染色体异常等出生缺陷的发生风险也相应增加。产前筛查是应对这类问题的重要措施。其中无创产前基因检测技术 (NIPT) 因具有安全，无创等优势，自发现以来迅速在唐氏综合症，爱德华氏综合症以及帕陶氏综合症的临床检验领域得到广泛应用。NIPT 通过采集孕妇外周血液中的胎儿游离 DNA 进行测序，并进行信息分析来获得胎儿的遗传信息，可用于识别胎儿常见的染色体变异。

1.2 问题重述

NIPT 的准确性主要依赖于胎儿性染色体浓度的判断。我们通常在孕妇孕期的 10-15 周之间对胎儿的性染色体浓度进行检测。男胎的 Y 染色体浓度达到或高于 4%，女胎的 x 染色体浓度无异常时，可认为 NIPT 结果较为准确。同时，为降低治疗风险，我们应尽早发现不健康的胎儿，及时进行干预治疗。

在实践中，孕妇孕周数以及其身体质量指数 (BMI) 均可能影响男胎 Y 染色体浓度，而孕妇的年龄，BMI，孕清等个体差异也会影响 NIPT 的准确性。因此，合理分组孕妇，并确定最佳检测时间点，对减少潜在风险至关重要。附件提供了某地区孕妇的 NIPT 数据，包括多次检测情况，测序失败情况，孕妇的各项身体指标以及检测各项数据，用于研究合适的 NIPT 时点和检测准确性。

要求基于这些数据建立数学模型解决以下问题

1. 建立相应的关系模型，分析男胎 Y 染色体浓度与孕妇的孕周数和 BMI 等指标的相关特性，并检验其显著性。
2. 临床表明，影响胎儿 Y 染色体浓度最早达标时间的重要因素为男胎孕妇的 BMI。根据男胎孕妇的 BMI 进行合理分组确定每组的 BMI 区间和 NIPT 最佳时间点，以最小化孕妇潜在风险，并分析检测误差对结果的影响。
3. 除 BMI 外，孕妇的身高，体重，年龄等均会影响胎儿 Y 染色体浓度达标时间。综合考虑孕妇身体状态，检测误差和 Y 染色体浓度达标比例，基于男胎孕妇的 BMI 进行分组，确定每组的最佳 NIPT 时点，以最小化孕妇潜在风险，并分析检测误差对结果的影响。
4. 女胎不含有 Y 染色体。综合考虑 x 染色体及 21 号，18 号，13 号染色体的 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 1 分析

企业通过抽样检测的方式判断一批零件的次品率是否小于标称值相当于统计学中的假设检验问题。可使用简单随机抽样抽取检测样本，由于企业需要做检测成本与检测效果之间的权衡，需找寻最佳样本量。可利用检测单价与检测样本量的乘积表现检测成本，利用假设检验的功效函数表现假设检验的效果，将二者线性组合为决策函数，寻找使得决策函数取极值的检测样本量作为最佳样本量。

2.2 问题 2 分析

企业需要解决下述四个决策问题：

- 零配件 1 是否检测
- 零配件 2 是否检测
- 成品是否检测
- 不合格成品是否拆解。

故共存在 16 种决策方案。定义最佳方案为利润期望最大的方案，利用二项分布和几何分布的性质，分别严格求出 16 种决策方案每生产出一个合格单品利润的期望。再将六种情型的具体数值代入利润期望公式，得出每种情形的最佳决策方案及决策指标值。

2.3 问题 3 分析

企业需要解决下述五类决策问题：

- 零配件 1-8 是否检测
- 半成品是否检测
- 不合格半成品是否拆解
- 成品是否进行检测
- 不合格成品是否拆解。

基于题目给出的 8 个零配件和 3 个半成品，问题二使用的对每种决策方案求利润期望的方法对应的决策方案数量过多，且无法推广至 m 道工序、 n 个零配件的生产情况，故考虑将该问题转化为一个多阶段决策问题，建立状态-决策模型，并通过动态规划算法来求解最优策略及决策指标值。

2.4 问题 4 分析

为衡量次品率取值可能存在的“误差”，可利用贝叶斯估计，选用共轭先验分布 Beta 分布，得到参数的后验分布。重新解决问题二，可将第二问中求出的利润期望公式视作次品率的函数，在本问的假定下，次品率服从某一分布，可进一步计算在次品率的后验分布下利润的期望。比较 16 种决策方案的利润期望得到最佳决策方案。重新解决问题三，分析可得唯一的不同在于，在问题三中未检测的零件、半成品 (组成零件均合格)、成品 (半成品均合格) 是否为次品均服从次品率固定的伯努利分布，问题四引入了 Beta 分布来描述次品率的不确定性。通过推导，得出新的次品概率分布，并在此基础上进行最优决策。

三、 模型假设

为简化问题，便于模型建立与求解，本文做出以下合理假设：

- **假设 1：**所有决策为确定性决策，适用于生产线上所有相应节点。
- **假设 2：**退回成品拆解后，若无法确定相应零件或半成品是否为次品，则必须进行检测。

说明：假设生产过程中不涉及概率性问题，每一项决策 (如是否检测、是否拆解等) 都是确定的。这意味着在任何时候，检测、拆解等操作要么执行，要么不执行，并且这一决策适用于所有相关节点。这样处理降低了流水线决策成本，符合实际生产中工厂流水线的情况。

说明：假设在拆解后的零件和半成品存在质量不确定性时 (即之前未对该半成品或零件进行检测)，必须进行检测。此假设考虑到企业质量控制的需要，确保每个生产阶段的次品得以及时识别和处理，更重要的是显著提升效率，防止次品在流水线上无限循环，合理地简化了模型。

四、 符号说明

为方便阅读，本文使用的主要符号及其说明如表 1 所示。

表 1 符号说明表

符号	说明
n	抽样检测样本量
v	售卖成品的单位利润
P	成品市场售价
C	合格成品的单位成本
St	t 时刻零件 1 – 8、半成品 1 – 3 及成品的状态变量
R	成品收益，包括销售收益和相应拆解后半成品或零件的回收收益
α, β	先验分布中的超参数

五、 问题一的模型建立与求解

5.1 模型建立

在频率学派假设检验的理论体系下，原假设和备择假设的地位存在差异，本问题中企业对于拒收与接收的态度也存在差异，故可以利用假设检验的思想处理该问题 [1]。而 (1) 和 (2) 两个情形下，不仅仅只有显著性水平的差别，原假设与备择假设互相相反，故在下文中分开对两个情景进行讨论。

引理 1 这是一个引理...

定义 1 这是一个定义...

$$E = mc^2$$

$$1 - \Phi \left(\frac{0.3u_{0.95} + 0.1 - p}{\sqrt{\frac{p(1-p)}{n}}} \right)$$

这是一个需要引用的重要公式，见公式 (1)。

$$E = mc^2 \tag{1}$$

$$w_1 Cn - w_2 \int_0^1 f(p) K(n, p) dp$$

下图 (图 1) 展示了一个示例。

这句话引用了文献[?]。现在所有引用都将自动变为上标。

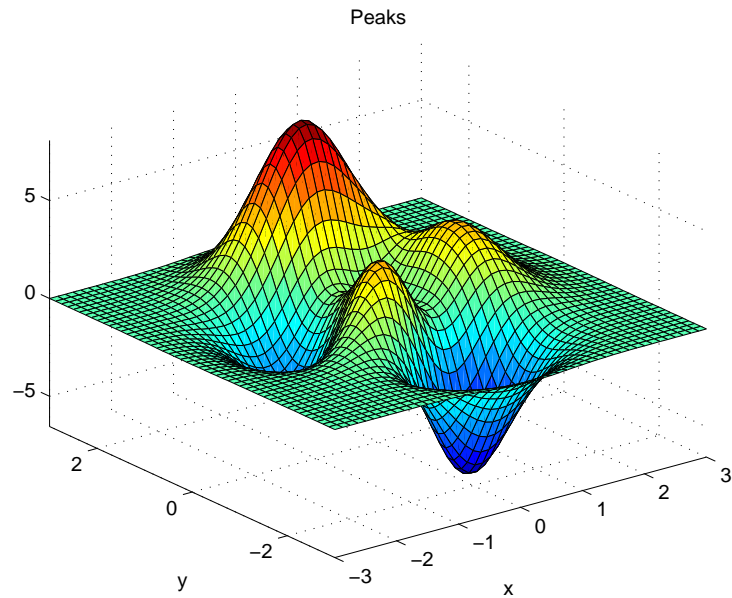


图 1 单图示例

5.2 模型求解

为求解上述模型，我们设计了如下步骤：

Step 1: 数据预处理...

Step 2: 利用 MATLAB/Python 实现... 算法...

Step 3: 结果后处理与可视化...

5.3 求解结果与分析

...

六、 问题二的模型建立与求解

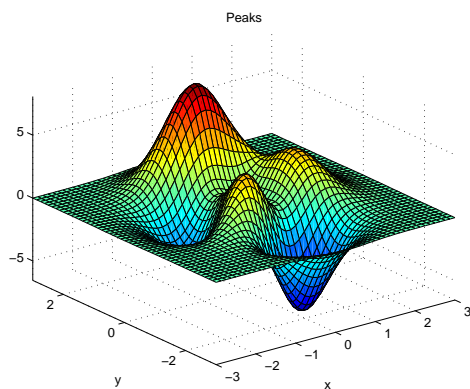
6.1 模型建立

对于问题二，我们考虑... 因素，建立了... 模型。如图 2所示，其中图 2a展示了...，图 2b展示了...。

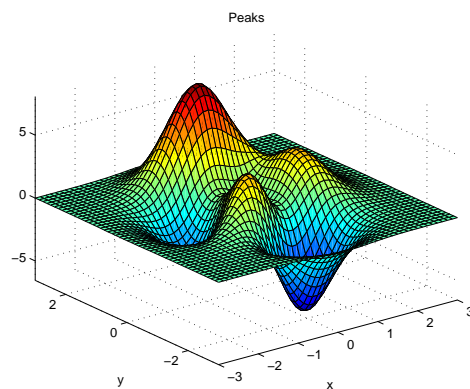
七、 模型的分析与检验

7.1 灵敏度分析

...



(a) 子图 A 的标题



(b) 子图 B 的标题

图 2 双子图示例

7.2 误差分析

...

八、 模型的评价与推广

8.1 模型的优点

- 优点 1: 模型结构简洁, 物理意义明确。
- 优点 2: 算法效率高, 求解速度快。
- 优点 3: 模型具有较好的通用性, 可推广至...

8.2 模型的缺点

- 缺点 1: 假设条件较强, 在... 情况下可能失效。
- 缺点 2: 未考虑... 因素, 与实际情况存在一定偏差。

参考文献

A 支撑材料列表

表 2 附录文件列表

文件名	功能描述
q1.m	问题一核心 MATLAB 程序代码
q2.py	问题二 Python 求解程序
data.xlsx	本文使用的原始数据与处理结果

B 核心代码

问题一 MATLAB 代码 (q1.m)

```
1 disp("Hello World!")
```

问题二 Python 代码 (q2.py)

```
1 print("Hello World!")
```