基于生产过程中检测方法的研究与优化决策*

刘昊 张英婷 陈佳 韦铸娥**

南宁学院信息工程学院，南宁 530200

摘 要 为帮助生产者更好地分析和决策生产过程中的零件检测问题，研究综合运用帕斯卡分布、单侧假设检验、二项分布、高斯分布及 Fisher 检验等统计方法，构建决策树模型与动态规划算法，并利用 Python 软件进行分析求解。该方法能有效优化检测决策，提高生产效率，并在保证产品质量的同时降低检测成本。实验数据表明，相较于传统检测方法，本文的优化策略在决策准确性和经济效益方面均有显著提升，适用于不同规模的数据集，并能适应多种工业生产环境。通过优化检测流程，企业可在控制成本的同时提升质量管理水平。

关键字 帕斯卡分布，单侧假设检验，高斯分布，Fisher 检验，动态规划

Research and Optimization of Inspection Methods in the Production Process

Liu hao Zhang Yingting Chen Jia Wei Zhue

College of Information Engineering of Nanning University
Nanning 530200 China
361762849@qq.com

Abstract—To assist producers in better analyzing and making decisions regarding part inspection during the production process, this paper integrates statistical methods such as Pascal distribution, one-sided hypothesis testing, binomial distribution, Gaussian distribution, and Fisher's test. Decision tree models and dynamic programming algorithms are developed, and Python software is utilized for analysis and solution. This effectively optimizes inspection decisions, enhances production efficiency, and reduces inspection costs while ensuring product quality. Experimental data show that compared to traditional inspection methods, the proposed optimization strategy significantly improves decision accuracy and economic benefits, making it suitable for datasets of different scales and various industrial production environments. By optimizing the inspection process, enterprises can improve quality management while controlling costs.

Keywords—Pascal distribution, one-sided hypothesis test, Gaussian distribution, Fisher's test, Dynamic Programming

1 引 言

在现代制造业中，产品质量控制是企业竞争力的核心要素之一。尽管我国已成为机械制造大国，但在国际市场上仍存在产品质量不稳定、生产效率偏低等问题[1]。质量检测是确保产品合格率、提高市场竞争力的重要手段之一。然而，传统检测方法存在着检测成本高、决策效率低、过度检测或检测不足等问题。因此，如何优化检测策略，在控制成本的同时确保质量稳定，已成为工业生产管理的重要研究方向[2]。

本文基于数学统计方法和优化算法，提出了一种改进的检测决策方案。通过构建数学模型，利用统计检验方法分析零件质量，并结合动态规划优化检测策略，以实现质量控制与成本优化的平衡。研究不仅有助于提升检测决策的科学性，还能提高企业的经济效益，为制造业提供更优的质量管理方案。

对于零部件的检测与决策问题，做以下几个假设：零件的质量是随机分布的；每一个零件被抽中的概率相等；抽样过程中不会有新的零配件加入；假设每个样品都具有相同的代表性。

根据已有的数据和抽样方法，分别求解以下问题：

问题1：分析不同的抽样检测方法的关系，依据零配件的标称值寻找接受和拒绝零配件的临界值，并给出最优的解决方案。

问题2：依据企业在生产过程中遇到的情况，合理判断零配件1、零配件2和成品是否进行检测做出最优决策。

问题3：企业在生产中遇到m道工序、n个零配件，对其零配件、半成品和成品是否进行检测做出一个最

*基金资助：南宁学院横向科研项(2024HX137, 2024HX146)
**通讯作者：韦铸娥 361762849@qq.com

优的决策方案，并给出相应指标。

问题4：假设问题2和3中零配件、半成品和成品的次品率均是通过抽样的方法得到的。调查结果相对于总体真实值的精确度与样本容量直接相关[3]。但在实际生产过程中，并非能知道某批零配件的次品率，基于问题1预测出该零配件的次品率，重新完成问题2、问题3。

2 模型的建立与求解

2.1 问题 1 的模型与求解

因为帕斯卡分布是描述在一系列独立的伯努利试验中，需要进行多少次试验才能达到固定次数的成功[4]。在问题1中，已确定了次品率，只需确定进行的试验次数即可。可通过帕斯卡分布，在次品率范围内，以不同的信度确定样品种所存在的次品零部件，以寻找寻预计检测的样本量。

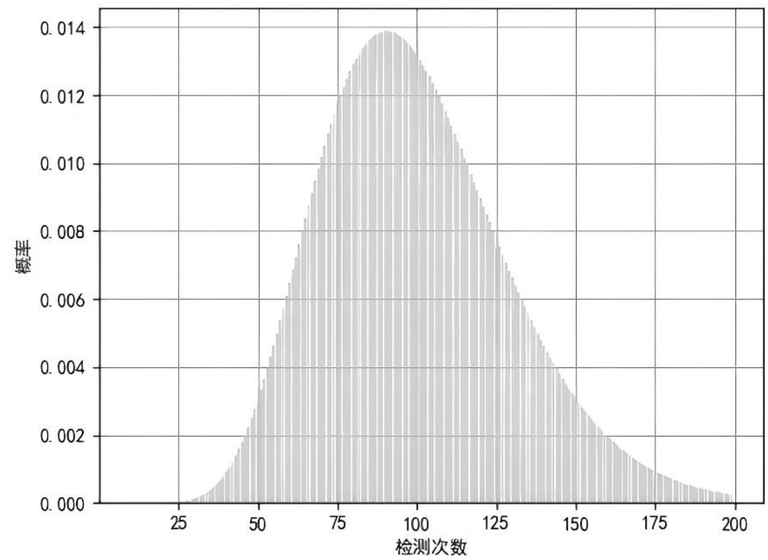

图 1 帕斯卡分布检测

由图1可看出，检测次数与次品出现的次数近似成正态分布。设定标称值为10%，运用帕斯卡分布原理计算可得，在次品率为10.6%的情况下，预计找到10个次品需要检测94.34个产品。这是基于次品率10.6%和找到10个次品的目标计算的平均检测次数，它不依赖于总体样本量，而是和次品率以及找到目标次品数有关。无论样本量多大，在每次检测时每个产品为次品的概率依然是10.6%。负二项分布计算的期望值是基于每次检测是否为次品的独立性和次品率的稳定性来得出的。

问题1还需要确定临界值k的值为多少时选择拒收和接收，这里引入累积分布函数，进行完整描述一个实随机变量X的概率分布，在$x \geq k$或$x \leq k$时的概率，其中：

$$P(x \geq k) = 1 - \sum_{i=0}^{k-1} P(x = i),$$

$$P(x \leq k) = \sum_{i=0}^{k} P(x = i)。$$

当信度在95%时，$P(x \geq k_r | p_0 = 0.1) \leq 0.05$,当信度在90%时,$P(x \leq k_a | p_0 = 0.1) \geq 0.9$，由上述两中模型联立解得方程的临界值$k_a$和$k_r$。

在实际生产中，由于样本量的数据过大，其他的前提条件保持不变条件下，可考虑使用高斯正态分布进行近似拟合二项分布。高斯分布广泛应用于统计学、化学、物理学、金融学等诸多领域[5]。分析本文研究可以发现，如果把许多小作用加起来看做一个变量，理论上可以证明这个变量服从正态分布，采样分布均值是近似地正态的，即使被采样的样本的原始群体分布并不服从正态分布[6]。正态分布信息熵在所有已知均值及方差的分布中最大，这使得它作为一种均值以及方差已知的分布的自然选择[7]。

通过使用阶梯图来表示整个样品数与接受的、拒接的临界值范围，阶梯图的具体绘图步骤如下：

Step (1) 设置标称次品率、显著水平（拒绝区域）和反显著性水平（接受区域）。

Step (2) 计算不同样本量下的临界值。

Step (3) 设置样品最大值，规定样品上限。

Step (4) 根据5.1.1中的结果对其图像进行验证。

利用Python绘制出阶梯图如图2所示。

从图2可明显看出，随着样本数N的增加，拒绝临界值和接收临界值均呈现上升趋势，这与数据的统计显著性逐渐增强有关，因为在大样本条件下，较小的

差异更容易被检测到。图2中的实线为拒绝临界值曲线，随着样本数的增加呈现出阶梯状上升，表明其临界值随着样本数的变化而不断提高。这意味着，随着样本量增大，对某种假设（如假设检验中的零假设）的拒绝标准越来越严格。

图2中的虚线为接收临界值曲线，虽然同样随着样本数增加而上升，但其增长速度明显慢于拒绝临界值。这说明在样本量增加的情况下，接收假设的标准的增长幅度远小于拒绝假设的临界值。这意味着即使在较大样本量下，接受假设的标准仍然比较宽松。这种差异导致随着样本数的增加，接收临界值与拒绝临界值之间的差距逐渐拉大。这种临界值差距的扩大表明，随着样本量的增大，更加倾向于拒绝一些不显著的假设（即假设的临界值越来越严格），同时对接受假设的判断标准变化较少。这反映了在大样本情况下，能更精确区分出显著性差异，因此对假设的判断更严格。

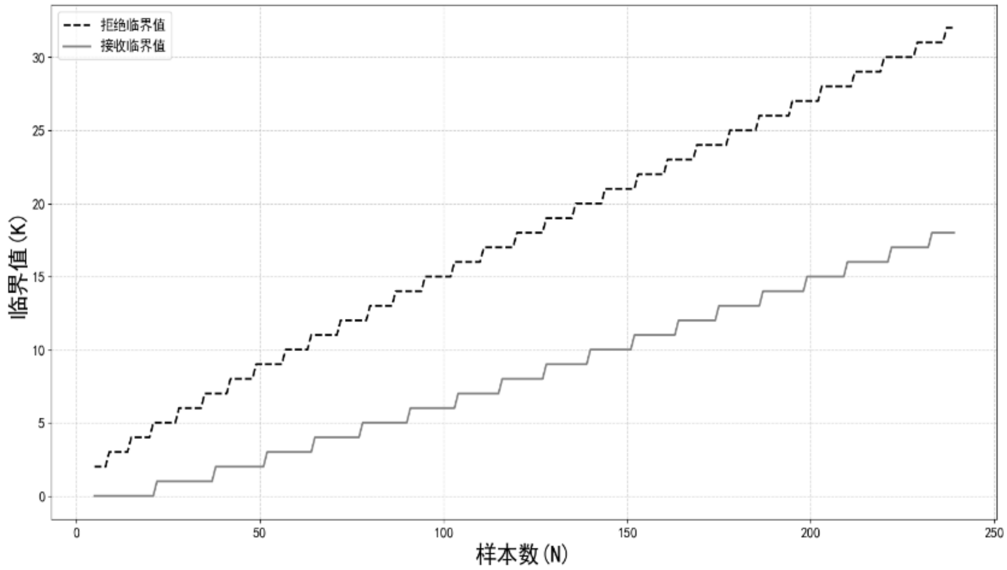

图 2 K-N关系图

2.2 问题 2 的模型与求解

首先要根据生产中遇到的情况进行不同零配件和成品是否检验的决策分析。由于假定生产中的次品率都是一定的，因此可直接采用决策树法搭配动态规划的方法进行所有组合的迭代计算。由于每个零配件或成品都可以自由选择是否检验，故可极端地认为所有零配件和成品全部检验或者所有零配件和成品全部不检验。然后在求解此问题上选择将每个零配件1、零配件2和成品的是否检验作为一个决策点，在对其进行动态规划处理，搭配逻辑如图3所示。

由检验与否以及是否拆卸组装易知，对于零配件1、零配件2和成品三个数关于是否检验的排列一个有2^3=8个结果，分别为：

(1) (False, False, False)；

(2) (True, False, False)；

(3) (False, True, False)；

(4) (False, False, True)；

(5) (True, True, False)；

(6) (True, False, True)；

(7) (False, True, True)；

(8) (True, True, True)。

其中，False表示不进行检测，True表示进行检测。

通过对每组的期望利润和期望成本进行计算，然后遍历每组的期望利润进行组间比较，从而寻找全局最优解。利用Python求解动态规划，针对具有代表结果的情况1和情况5，运算结果如表1、表2所示。

针对两种情况，动态规划进行迭代计算得出的结果，进行Python的数据遍历，看出针对情况1 和情况5的决策方案分别为：全部不检测和仅零配件2检测，由已知数据表可知，针对情况1，由于情况1的次品率较低，但是检测成本相对于调换成本略高，故如果进行检验的话检测到次品的概率较低，即使检测出相较于调换费用，检测费用仍是高出。针对情况5，零配件2的次品率要普遍高于零配件1和成品，且零配件2的检测成本要远远小于调换成本故选择对零配件2进行检验，通过遍历的迭代的结果可知，这种决策方案确实是最大利润。详情分析如下：

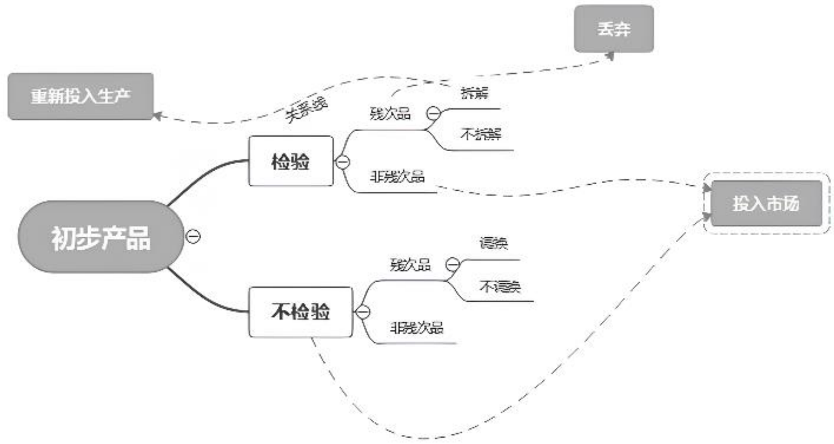

图 3 检验逻辑图

表 1 动态规划检测结果(情况1)

组合	期望利润(元)	期望总成本(元)
(False，False，False)	4530.00	510.00
(True，False，False)	4390.00	650.00
(False，True，False)	4290.00	750.00
(False，False，True)	4230.00	810.00
(True，True，False)	4150.00	890.00
(True，False，True)	4090.00	950.00
(False，True，True)	3990.00	1050.00
(True，True，True)	3850.00	1190.00
最佳方案	最大期望利润	
（False，False，False）	4530.00	

注： 组合中的True为检验，False为不检验

表 2 动态规划检测结果（情况5）

组合	期望利润(元)	期望总成本(元)
(False，False，False)	4210.00	830.00
(True，False，False)	3470.00	1570.00
(Fals，True，False)	4230.00	810.00
(False，False，True)	4050.00	990.00
(True，True，False)	3490.00	1550.00
(True，False，True)	3310.00	1730.00
(False，True，True)	4070.00	970.00
(True，True，True)	3330.00	1710.00
最佳方案	最大期望利润	
(False, True, False)	4230.00	

注： 组合中的True为检验，False为不检验

首先，针对情况1，由于该情况中零配件的次品率相对较低，且根据企业在生产中遇到的情况，检测成本与调换成本之间的关系是关键影响因素。在情况1中，虽然存在一定概率检测到次品，但由于检测成本明显高于调换成本，即便检测到次品的可能性较低，进行检验所产生的费用仍然大于直接调换的成本。因此，企业在这种情况下选择不进行任何检测，以避免不必要的成本浪费。这种决策不仅符合成本最小化的原则，也确保了企业资源的有效利用。

其次，针对情况5，零配件2的次品率显著高于零配件1和成品，这意味着该零配件在生产过程中更容易出现问题。因此，检测零配件2成为一个重要的控制环节。同时，零配件2的检测成本远低于其可能带来的调换成本，这为决策提供了关键依据。通过检测零配件2，企业能够有效控制生产中的质量风险，而不必承担更高的调换费用。

通过上述方法的重复操作，便可得到所有假设情况的结果如表3所示。

表 3 决策方案

	零配件1	零配件2	成品
情况1	不检验	不检验	不检验
情况2	不检验	不检验	不检验
情况3	不检验	不检验	不检验
情况4	检验	不检验	检验
情况5	不检验	检验	不检验
情况6	不检验	不检验	不检验

2.3 问题 3 的模型与求解

问题3在问题2的基础上进行建立，将原有的问题2中的零配件1、零配件2的零配件扩展为n个，将原有的一到加工工序拓展至m个，由于该零配件数量呈现指数型增长，故问题3相比问题2有更为复杂的程序联系与计算过程。

至此，问题2已经完全解决。

为了有效应对这些新增的变量与复杂性，可以继续使用问题2中的思路与方法。首先，针对多零配件和多工序的情境，继续采用与问题2相似的动态规划与决策树相结合的方式，通过逐步细化决策点，对整

个生产流程进行更加精细的划分。在这一过程中，在问题2的基础上进行了进一步的改进和优化，不仅扩展了模型的适用范围，还增强了算法的适应性，以便在处理n个零配件和m道工序时，仍能够快速准确地找到最优的决策方案。

具体来说，首先将零配件的扩展问题分解为多个子问题，每个零配件的检测方案、检验成本、次品率等因素都会独立考虑并通过动态规划进行迭代计算。同时，对于多道工序的扩展，也进行了详细的工序划分，确保每一道工序在不同条件下都能得到最优的处理与决策。通过引入多层次、多维度的优化模型，不仅保留了问题2中的解决思路，还在问题3的复杂背景下进行了有效的扩展。

由于不同的零配件的次品率相同，但其购买单价和检测成本不同，故对不同的零配件是否检验将对最终的利润结果产生影响。同时半成品和成品的检测成本较高，故有关产品与半成品是否进行检查、检测后是否拆解等问题，对最终的期望利润会产生较大的影响。

动态规划是运筹学的一个分支，是解决多阶段决策过程最优化的一种数学方法，适用于解决最优路径问题、资源分配问题、生产调度问题、库存问题、装载问题、排序问题、生产过程最优控制问题等[8][9][10]。

动态规划首先需要进行递归关系的建立，即通过从大到小的方式进行问题拆解，从而递归出最优解，利用如下的递归关系式：

$$f(i,w) = \max(f(i-1,w), f(i-1,w-w_i)+v_i)$$

使用归纳步的方法进行求解，即：

$$f(s_n) = \min_k\{g\big(f(s_k)\big)+h(s_k)\}, k<n$$

或

$$f(s_n) = \max_k\{g\big(f(s_k)\big)+h(s_k)\}, k<n$$

为了开始递归计算，需要设置边界条件，即最小值问题的解。在问题3中，当没有物品可以放入时（即i=0），无论容量如何，最大价值都是0，因此：

在问题2的基础上，根据题目要求，继续完善程序支路，将每个零配件、半成品和成品的次品率，购买单价，检测成本进行参数自定义，根据上述的相关期望公式进行函数定义，利用动态规划模型遍历所有组合，计算出每种组合的相应利润。

为了能够更直观化的展示各组路径（每组组合路径），对布尔值的标题字符进行二元化处理，其中True为0，False为1，展示部分结果如图4所示。

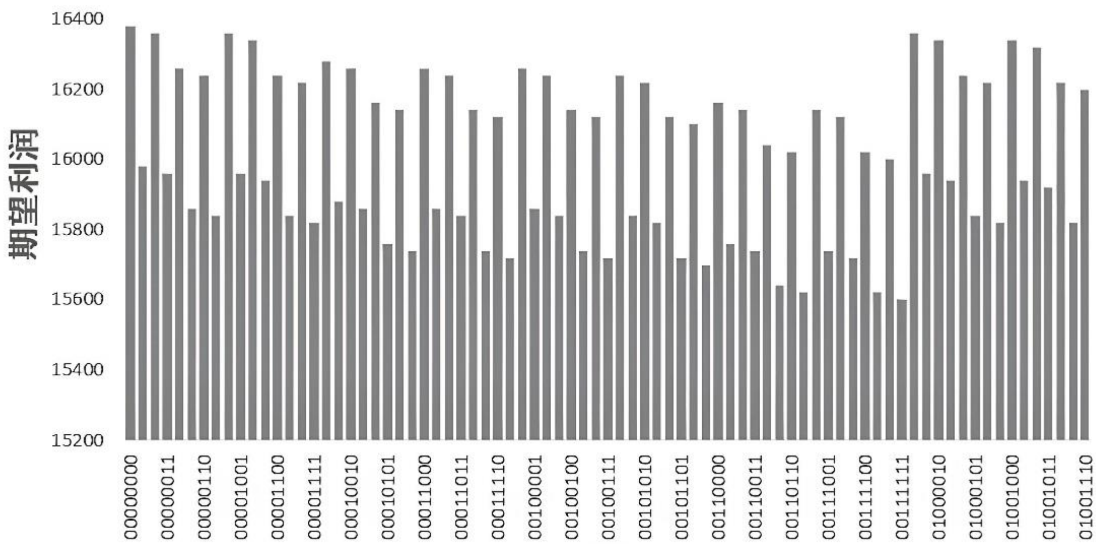

图 4　期望利润与检验方式

由图4可直观看出，多数高柱对应的检验方案均为检测成品和半成品的不检验，因为成品和半成品的次品率较低，故出现次品的可能性较小。又因为成品和半成品在生产过程中出现次品的可能性较小，因此进行检测的必要性相对较弱。进一步分析可知，成品和半成品的检验成本与装配过程中产生的成本相比要高出不少，尤其是当零部件的次品率较低时，检验的成本往往超出可能带来的经济效益。此外，成品和半成品的拆解成本和调换成本相对较低，使得在成品或半成品出现问题时，选择进行调换反而可避免不必要的检验成本。由于其检验成本与装配成本要高于拆解成本，所以对成品和半成品进行检验会产生较多的成本费用，从而使得整个生产流程中的总成本得到有效控制，趋势如图所示。即针对问题3，所用零配件与成品、变成品均不检验时，期望利润达到最大值。

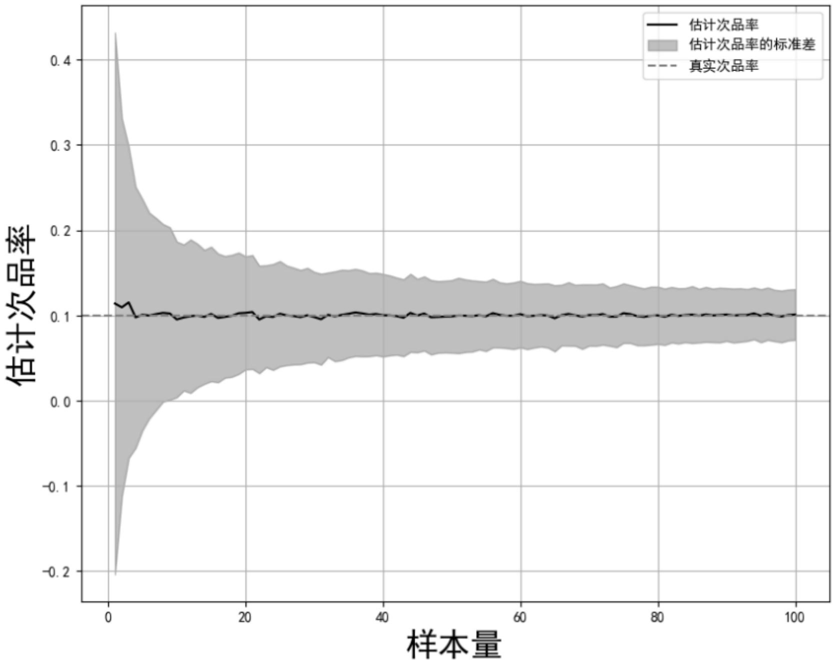

图 5 样本量与次品率

2.4 问题 4 的模型汇总与求解

要针对不同零配件、成品和半成品的检验方式做出的决策，通常没有确定的次品率。问题4考虑到此问题，将问题2和3中的相关次品率丢弃，参考问题1中次品率与标称值质之间存在的某种关联，对次品率进行预测。在预测次品率的基础上，问题4继续沿用问题2和3中的方法对生产流程进行分析与优化。预测的次品率数据成为了后续决策过程的核心依据，通过运用这些数据，重新审视并处理了问题2和问题3中的相关决策，最终给出了更加适应现实情况的具体解决方案。

基于问题1中的的相关条件，如标称值为10%，利用鲁棒优化的方法进行假设，将样本量划分在0-100之间，根据二项分布计算残次品的期望数和计算每个样本量下的次品率及其标准差，来预测次品率，利用Python进行绘图，如图5所示。

由图5可知，随着样本量的增加，零配件的次品率收敛与0.1附近。

针对问题2，将零配件和成品的次品率假定在0.1上下随机浮动，使用原问题2的方法，即使用决策树和动态规划进行求解，得到新求解的结果如表4所示。

上述方法可对不同的次品率找出最优的决策方案，对于问题4中的重复问题2将不做出假设次品率方案。

表 4 决策方案

组合	期望利润(元)	总成本(元)
(False, False, False)	4297.00	873.67
(True, False, False)	4259.00	911.67
(False, True, False)	4265.00	905.67
(False, False, True)	4281.00	889.67
(True, True, False)	4227.00	943.67
(True, False, True)	4243.00	927.67
(False, True, True)	4249.00	921.67
(True, True, True)	4211.00	959.67
最佳组合：	最大期望利润:	
(False, False, False)	4297.00	

注：（情况4）次品率估计：零配件1 = 0.10，零配件2 = 0.11，成品 = 0.08

针对问题3，思路与解法和上述解法相同，但是在确定次品率上需要进行微调整，由于问题3的零配件和加工工序较多，需要扩大次品率在0.1处的浮动范围，以避免重复，继续使用Python对所有的数据进行遍历比较，得出的部分结果如图6所示。

由图 6 不难看出，随着零配件的次品率改变，所呈现的决策方案相比较原问题 3 的图像有较大的差异，但其数据量呈现显著增加的区域还是关于成品与半成品这类调换费用和检测费用较大的一类。

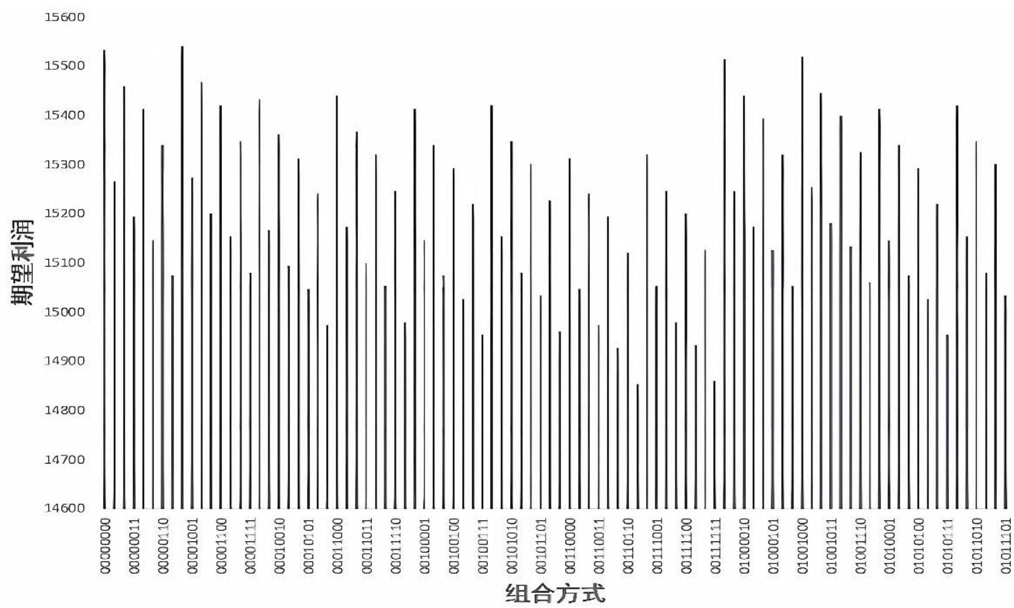

图 6　组合方式与期望利润关系

3　结束语

为解决大企业在生产过程中的检测问题，提出了正态分布近似、Fisher精确检验、期望成本模型和抽样检测等方法，能应对不同规模的数据和复杂的业务决策需求，有较强的灵活性与实用性。本文遍历组合和评估期望利润的方式、模型，帮助企业优化检测策略，实现利润最大化并控制成本。然而，随着检测组合、样本量和参数数量的增加，计算复杂度会迅速上升，可能影响大规模数据下的计算效率。此外，模型高度依赖于数据的准确性，尤其是次品率和检测成本等参数，若数据存在偏差，决策的准确性也将受到影响。因此在实际应用中，需平衡数据准确性与计算复杂度，才能找到最优解并为业务提供有力的支持。

参 考 文 献

[1] 赵淑琴.利用质量管理技术提高产品质量策略研究[J].中国设备工程,2019,(17):29-31.

[2] 侯先荣,吴奕湖.21世纪的新质量观[J].工业工程与管理,2001,(01):13-15.

[3] 邵志强.抽样调查中样本容量的确定方法[J].统计与决策,2012,(22):12-14

[4] 叶利娟.帕斯卡分布展示的教学思想[J].数学学习与研究,2012,(23):128+130.

[5] 刘爔.多维折叠正态分布的性质和参数估计[D].上海：上海师范大学,2024.

[6] 胡默迪.基于孤立森林的工控网络入侵防御系统的研究与应用[D].北京：北京邮电大学,2020

[7] 陈根方.中国工尺谱的数字实现研究[D].上海：上海大学,2011.

[8] 满孝虎.基于动态规划方法的瓦斯抽采钻孔施工设备的更新研究[J].中国设备工程,2024,(14):126-128.

[9] 褚一帆.工作流中若干问题的研究[D].南京：南京航空航天大学,2012.

[10] 赵娟,王建新.用动态规划方法求解最短运输路线问题[J].现代电子技术,2012,35(17):120-122