Optimizing Quality Control in the Food Industry: Maximum Likelihood Estimation

Eva Pérez S.

ABSTRACT

Quality control (QC) is a vital process for ensuring and enhancing product quality across various industries. The approaches to QC vary depending on the specific product or industry, going through different methods and techniques. In this context, this paper explores the potential of integrating Maximum Likelihood Estimation (MLE), a statistical method commonly used in Artificial Intelligence, to enhance the efficiency and accuracy of manual quality control inspections. The investigation seeks to ascertain whether such statistical techniques could offer greater utility in the future, given the exponential advancements in technology. This examination is particularly relevant to the food industry, where safety and quality standards are taken great care of, and where regulatory testing, inspections, and guidelines such as Good Manufacturing Practices (GMP) play an important role. The paper reinforces the need to adapt and look for new methodologies in response to evolving technologies, with a focus on their potential impact in industries like food production.

Keywords: Quality Control, Maximum Likelihood Estimation, Machine Learning, Artificial Intelligence

INTRODUCTION

Quality control is a process where a business looks for a way to ensure that product quality is being maintained or improved, QC methods and techniques may be different depending on the specific product, service, or industry in question. There is not just a single approach to quality control, and different businesses or sectors could employ various methods to measure and ensure quality. Hayes (2023) As an example, as we will go through in this paper, in the agrifood sector, this involves a combination of regulatory testing, inspections, and guidelines like GMP.

Taking into account all of the above, we will have a look to determine if a statistical method like Maximum likelihood estimation, used in Artificial Intelligence, can possibly enhance efficiency and reduce mistakes in a manual quality control inspection. The purpose of this investigation is to get to know if these kinds of methods could be more useful in the near future or if it's necessary to look for a different approach when it comes to the food industry, always keeping in mind the exponential improvement of new technologies and how it can make a big impact in such industries.

MAXIMUM LIKELIHOOD ESTIMATION

MLE generally speaking, is a statistical method used to estimate the parameters of a statistical model or probability distribution based on observed data. It works by finding the parameter values that make the observed data most probable or likely under the assumed model.(Myung, 2003)

As we can see in the next formula:

$$\hat{\theta}_{\mathrm{MLE}} = \arg\max_{\theta} \mathscr{L}(\theta \mid \mathbf{x})$$

MLE starts by picking a statistical model that fits the data well. This model has parameters (usually represented as θ) that we want to find.

Next, we use a likelihood function (L) to measure how well this model fits the actual data. The likelihood function, denoted as L(θ |data), depends on the parameters (θ) and the data we have.

To make the math easier, we often work with the log-likelihood function (LLF), which is the natural logarithm of L(θ |data), written as log(L(θ |data).

MLE's main job is to find the parameter values (θ) that make the log-likelihood as big as possible. This is done using math techniques like calculus or optimization algorithms.

The values of θ we find are called maximum likelihood estimates (θ_{MLE}). They represent the best-fit parameters for the chosen model, given the data.

In real-world applications, MLE doesn't just give us point estimates for θ. It also helps us understand how confident we can be in those estimates, usually by calculating confidence intervals or standard errors. This helps us to confirm how reliable our parameter estimates are. Firth (1993)

QUALITY CONTROL IN FOOD INDUSTRY

Quality control in the food industry is a critical process that ensures the safety, consistency, and overall quality of food products. It involves a wide range of activities, procedures, and standards to meet regulatory requirements and consumer expectations.Hayes (2023) Here are some of the main aspects of quality control in the food industry:

- Process Controls
- Regulatory Compliance
- Supplier Quality Management
- Hazard Analysis and Critical Control Points (HACCP)
- Allergen Management
- Employee Training
- Traceability
- Shelf-Life Testing
- Documented Quality Procedures
- Packaging and Labeling
- Consumer Feedback
- Continuous Improvement
- Quality Testing and Analysis

Each one represents some kind of safety in a certain area, but for MLE to be used properly we need a different type of approach and data recollection to support our methodology in the best way possible.

0.1 Data collection

To consider a method like MLE, it's necessary data that describe the variability or characteristics of the quality parameter of interest. In this section are described the data sources and types of quality control data used in this study while explaining the process of data collection and any steps that come with it:

Sample Data Data points that represent measurements or observations related to the quality parameter wanted to be estimated. These data points can come from samples taken at different points in the production process.

Parameter to Be Estimated You should specify the parameter you want to estimate using MLE. For example, if you are interested in estimating the mean (average) of a quality parameter, you would specify that as the parameter of interest.

Distribution Assumption You need to make an assumption about the probability distribution that describes the data. Common distributions used in MLE include the normal distribution, Poisson distribution, binomial distribution, etc. The choice of distribution depends on the nature of the data.

Likelihood Function Based on your distribution assumption, you need to formulate the likelihood function. The likelihood function quantifies the probability of observing your sample data given the parameter you want to estimate. It is the core component of MLE.

Log-Likelihood Function It's often more convenient to work with the log-likelihood function, which is the natural logarithm of the likelihood function. This is because it simplifies calculations.

Optimization Technique You would need an optimization technique to maximize the log-likelihood function with respect to the parameter of interest. Common optimization methods include numerical optimization algorithms like gradient descent, Newton's method, or built-in optimization functions in software packages like Python's SciPy library.

Initial Parameter Guess Many optimization algorithms require an initial guess for the parameter. You may need some prior knowledge or a reasonable guess to start the optimization process.

Sample Size The size of your sample data can significantly impact the accuracy of MLE estimates. Larger sample sizes generally provide more precise estimates.

Uncertainty Assessment After obtaining the MLE estimate, you should assess the uncertainty around the estimate. This often involves calculating confidence intervals or standard errors to calculate how confident you can be in the parameter estimate.

Quality Control Context You should have a clear understanding of how the parameter estimate fits into your quality control context. What action or decision will be based on this estimate, and what are the quality standards or tolerances you need to consider?(Escobar, 2018)

0.2 Good Manufacturing Practices

Part of regulatory compliance, food must follow strict regulations and standards set by the government (FDA in USA, EFSA in Europe). They include proper labeling, hygiene and proceed with the GMP. GMP establish the conditions and requirements necessary to ensure hygiene in the food chain and production.

GMP were developed by the **Codex Alimentarius** with the aim of protecting the customer. It includes several basic operating conditions and procedures that any food business must comply with. Some of them are:

- Construction and design of food premises.
- Training for employees.
- Proper maintenance of company equipment and tools.
- Correct use of chemicals, including chemical agents, pesticides and lubricants.

- Identification and storage of waste inside and outside the company.
- Cleaning of facilities, equipment, tools, floor, walls and ceiling.
- Effective pest control program.

All these related to the previous main aspects of QC, but with a proper follow-up and documentation that provides proper management in case something could be modified or improved for the better of the food industry.

(nqa, NA)

0.3 Previous studies

Microbiological Quality Control: MLE can be applied to estimate parameters related to microbiological quality in food products. For example, it can be used to estimate the mean and variance of microbial counts to assess food safety. Cypess (2012)

Sensory Evaluation: MLE can be used to estimate parameters related to sensory attributes of food products, such as taste, texture, and aroma, based on sensory evaluation data. Choi (2012)

Shelf-Life Estimation: MLE can help estimate the parameters of food deterioration models, allowing for the prediction of a food product's shelf life under different storage conditions. Bahk (2021)

Contaminant Detection: MLE can be applied to estimate parameters related to contaminant detection in food, such as the concentration of a specific contaminant in a batch of food.

Traceability and Recall Management: MLE can be used to estimate parameters related to batch-to-batch variability and traceability in the food supply chain, aiding in recall management in case of quality issues.

Process Optimization: MLE can be applied to optimize food processing parameters to achieve desired product quality attributes while minimizing costs and waste.

Product Authentication: MLE can be used in combination with authentication techniques (e.g., DNA analysis) to estimate parameters related to product authenticity and origin in cases of food fraud.

Consumer Preference Analysis: MLE can assist in estimating parameters related to consumer preference models, helping food companies tailor their products to consumer tastes. (Guo, 2020)

In essence, MLE is a versatile statistical tool that can be applied in quality control to estimate parameters essential for maintaining and enhancing product quality, safety, and consistency in various industries. It helps decision-makers make informed choices based on data-driven insights.

RESULTS

Incorporating Maximum Likelihood Estimation (MLE) into the food industry's quality control processes can offer several benefits:

Improved Parameter Estimation: MLE can provide more accurate estimates of key parameters in statistical models used for quality control. For example, it can help determine the mean and variance of product attributes, which are critical for maintaining consistent quality.

Enhanced Process Monitoring: Can be applied to monitor critical quality parameters in real-time. By continuously updating estimates based on incoming data, it enables quick detection of deviations from quality standards, allowing for immediate corrective actions.

Reduced Sampling Costs: Can help optimize sampling plans by determining the minimum number of samples required to achieve a desired level of confidence in quality control. This can lead to cost savings in terms of materials and labor.

Early Defect Detection: When integrated with data analytics and machine learning, can predict and detect potential quality issues early in the production process. This allows for proactive measures to prevent defects before they occur.

Process Optimization: By accurately estimating process parameters, MLE can aid in process optimization. This can result in higher yields, reduced waste, and improved overall product quality.

Traceability and Recall Management: Can facilitate better traceability by providing detailed information about product attributes and process conditions. In the event of a recall, this information can help pinpoint affected products quickly.

Customized Quality Standards: Allows for more precise tailoring of quality control standards based on historical data. This enables businesses to set quality thresholds that are specific to their processes and product variations.

Data-Driven Decision-Making: Empowers decision-makers with statistically rigorous insights. It enables them to make informed choices regarding quality control strategies, process adjustments, and resource allocation.

Regulatory Compliance: Many food industry regulations require statistical methods for quality control. MLE provides a compliant and scientifically rigorous approach to meeting these requirements.

Continuous Improvement: By continuously updating parameter estimates as new data becomes available, MLE fosters a culture of continuous improvement in quality control practices. It allows businesses to adapt to changing conditions and consumer preferences.

In summary, integrating MLE into quality control processes in the food industry can lead to more efficient, data-driven, and precise quality assurance. It helps businesses maintain high product quality, reduce costs, and stay competitive in a rapidly evolving industry.

Discussion

There are some implications while considering how effective would be in a real-life scenario, for example in a big industry it would facilitate their data processing as it would require less work labor, but in some cases, small industries wouldn't have the same adaptability as it would require data they don't have yet or it's not sufficient to adapt the AI.

Talking about resources, without including costs, it depends on how large the industry is using it and how it would be implied, it's not the same recollecting data and the processing in *reducing sampling costs* as in *early defect detection*. Both of them can be using the MLE statistical method, but the information they collect is different from one another.

Looking at the Data collection in section 0.1, it's enough context to assume that when isn't possible to get this type of data, MLE wouldn't have the same impact as it should normally, is a must for a good recollection of data to have a sufficiently good trained algorithm to help, rather than a poor recollection of data and a mediocre algorithm who needs to be checked manually to improve.

CONCLUSION

While MLE and AI offer significant advantages for quality control, their effectiveness depends on factors like data quality, the quality of algorithms and models, and the expertise of the personnel implementing and managing these technologies. Therefore, these technologies may not be suitable for all industries or processes, and a careful assessment of their applicability is necessary.

MLE and AI have the potential to enhance quality control by providing more advanced data analysis, predictive capabilities, and automation. However, their success depends on how well they are integrated into existing quality control processes and the specific needs of the business or industry.

There are a lot of industries, not only food-related, that are using these techniques, the most popular one, *Microbiological Quality Control*, can be as a reference as they have been using this technique for more than 20 years.

Looking forward to the near future, in large industries, it has more areas of implementation and it would absolutely be of use when more resources are needed, there's no guarantee it will always be right, it all depends on the trained algorithm and the data used.

GMP Good Manufacturing Practices

QC Quality Control

MLE Maximum Likelihood Estimation

REFERENCES

Bahk, G. J. (2021). Microbial-maximum likelihood estimation tool for microbial quantification in food from left-censored data using maximum likelihood estimation for microbial risk assessment. *frontiers*.

Choi, S. E. (2012). Sensory evaluation. *Johnes and Brlett Learning LLC*.

Cypess, R. H. (2012). Microbiological quality control. *ATCC*.

Escobar, C. A. (2018). Machine learning techniques for quality control in high conformance manufacturing environment. *Sage Journals*.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika Trust*.

Guo, M. (2020). Consumer preference analysis: A data-driven multiple criteria approach integrating online information. *Science Direct*.

Hayes, A. (2023). Quality control: What it is, how it works, and qc careers. *Investopedia*.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*.

nqa (NA). Gmp: Gestión de la inocuidad de los alimentos. *ORGANISMO DE CERTIFICACIÓN GLOBAL*.