**E_EOR1_ACSK** **VU Amsterdam**

2023-2024

# PI Game June 12, 2024

## 1 Game description

Welcome to the 2024 Annual PI ($\pi$, Probability and Inference) Game! In this game, you will have the chance to put your course-acquired knowledge and skills into practice. We will delve into a dataset from a local bakery chain store in Vilnius, Lithuania, known for its diverse assortment of bread and desserts, particularly its traditional Lithuanian pies. To reduce food waste, the bakery has partnered closely with a local feed mill. Surplus food is purchased by the mill at a fixed price, but the bakery is responsible for organizing transportation. Our analysis will focus on one of the store's bestselling items, sold by weight. For further context, additional background information can be found in Section 2.3. Our main goal is to provide valuable recommendations to bakery store managers regarding the optimal quantity of traditional food to prepare both in the short and long term. This, in turn, aims to help minimize losses and improve operational efficiency.

- Deadline: *June 28 at 6 am.* Feel free to submit your project report ahead of schedule if you prefer.
- Dataset: BakeryData2024_Vilnius.xlsx.
- $\pi$-Awards: As stated in the course manual, you will receive bonus points if your team receives the $\pi$-award. We will select the top five teams based on the quality and creativity of their written reports and award them bonus points. The team members who secure first place will receive an additional 0.5 points added to their raw grades, while second place will receive 0.4 points and third to fifth place will receive 0.3 points.
- The tasks comprise two main components. All teams must complete the first part, consisting of the primary tasks outlined in Section 2. For teams seeking extra challenges, a bonus point task is available in Section 3.

  It is important to note that your answer to the bonus point task will only be assessed if you attain a score of 80 or higher (out of 100) in the main tasks described in Section 2.
- *Strict* format requirements:
  (a) The minimum font size required is 11pt. The margins are not a concern for us.

(b) The entire assignment, including the title, group members' names, and reference list, should not exceed 16 A4 pages. There are no strict page limits for the responses to the bonus point task, but it is recommended to keep it within a maximum of 10 pages.

(c) As stated in the course manual (Section 4.1), if a student is reported by their group mates for lack of involvement, we will notify the exam office.

Please append a separate page at the end of your report, where each group member provides a brief statement detailing their contribution. This statement will not be included in the 16-page limit. Here is an example of such a statement: A and B conceived the idea presented in the report. A developed the theoretical framework and conducted the computations. C and D validated the analytical methods. B provided valuable insights and encouraged A to explore a specific aspect. All authors actively participated in the discussions, reviewed the results, and contributed to the final manuscript, except E.

(d) Please submit your work through Canvas Assignments. Ensure that you include:

- A written report solely in PDF format.
- A zip file containing your Python codes to replicate your results.

- The writing assignment will be graded as pass or fail, primarily based on Task 1 (p. 3), alongside an overall evaluation of the remaining answers. To pass, it is crucial to integrate the feedback received on the writing assignment adequately into the final report.

  - Linguistic accuracy and range: Your text should have only a few errors in grammar, vocabulary, punctuation, and spelling. Any mistakes should not hinder comprehension. Furthermore, your writing should demonstrate a wide range of vocabulary choices and grammatical structures.
  - Academic style: Your text should adhere to an appropriate discipline-specific academic style, which includes the use of formal language, an objective perspective, and the correct academic idiom.
  - Text structure: The overall coherence of your text is essential, with well-structured body paragraphs that focus on a single theme and feature clear topic sentences.
  - Clarity of formulation: Your sentences should be logically ordered, with clear connections between different ideas. Transitions between sentences should be appropriate and effectively convey the intended meaning.

- We recommend structuring the report by addressing the questions sequentially. Reports (and presentations) prepared using LaTeX are highly encouraged. A report template specifically for $\pi$-game is available on Canvas in the "LaTeX" module for your convenience.

- Read through your report before submission. Do not hand it in without reviewing and polishing it yourselves. It must be concise without redundant sentences.
- If you wish, choose an impressive team name.
- To successfully complete the course, each group member is required to present a portion of the project results on *June 28*. The details and schedule for the presentation day will be announced in Week 3.
- A gentle reminder: Engaging in solving tutorial exercises during practical sessions can greatly benefit your progress in the $\pi$-game. Neglecting these practical exercises may present additional challenges during the project.

# 2 Main tasks

The final report should cover the following four main task categories. Section 2.1 provides guidance on writing an introduction to the newsvendor problem. Section 2.2 consists of three theoretical tasks aimed at addressing questions in the practical section, without necessitating the use of the dataset. Section 2.3 comprises three practical tasks related to the Vilnius bakery data. In Section 2.4, a critical review of your data analysis is requested. The bonus point task mentioned above will be detailed in Section 3.

The tasks marked with asterisks (*) are more challenging compared to the others, with the number of asterisks correlating with the level of challenge. More asterisks signify greater difficulty.

## 2.1 Writing an introduction

The task in this section mirrors the process of writing the introduction section of a research article or thesis, similar to Section 1 in Beutner et al. (2023) and Friedrich and Lin (2024). In the introduction, a broad discussion on the research question's background is typically provided, as demonstrated in the opening paragraph of the aforementioned papers. This task aligns with the writing assignment, where you have previously conducted an initial study on the newsvendor problem.

**Task 1 (Pass/Fail)**

Building upon your writing assignment, please address the following points. Feel free to delve into comprehensive and relevant discussions. Your response should, nevertheless, not exceed two A4 pages in total.

(a) Begin by describing and summarizing the newsvendor problem (NVP).

(b) Motivate the importance of studying the NVP by presenting some interesting real-life examples. Explore applications related to the NVP and explain the reasons behind their relevance.

(c) Discuss the primary challenges in the real-life examples you have researched. Additionally, explore the methods introduced in this course that can be applied to tackle these challenges.

(d) Explore the applicability of these methods to other studies beyond the NVP. Offer at least one example of such studies and discuss their connection to the existing literature.

(e) Transition smoothly from the broad discussion above to the specific problem that will be investigated, focusing on the Vilnius bakery data. Introduce and motivate this specific problem.

(f) Determine which methods introduced in this course will be applied to address this specific real-life problem. Highlight the aspects in which these methods are beneficial for solving the problem. If possible, establish connections to the previous discussion in points (c) and (d).

(g) Summarize the main findings obtained from addressing this specific problem and transition to the remaining questions outlined in Sections 2.2 to 2.4, as well as Section 3 if you have completed the bonus point task.

While this task is presented as the first one, it is recommended to tackle it after completing the other tasks in the subsequent sections, particularly when addressing parts (f) and (g). Furthermore, the suggested order of addressing points (a) to (g) is purely advisory and not compulsory. You are encouraged to address these points in an order that best aligns with your discussions.

### 2.2 Theoretical tasks

In econometrics research, empirical motivation is crucial when writing an article or thesis. With this motivation in mind, it is customary to propose a general model and method before addressing practical questions. An example can be found in Section 3 of Friedrich and Lin (2024), where theoretical results are discussed prior to exploring empirical problems in their Section 6.

Similarly, in this task category, we extend the profit function, as taught in the lectures, to incorporate additional factors such as the clearance price and the shipping cost for unsold goods. We then delve into the theoretical properties of parametric and nonparametric methods, utilizing simulations. These theoretical properties offer valuable insights into the practical tasks at hand. Importantly, the use of the Vilnius bakery data is not necessary in this section. By following this approach, we aim to provide a solid foundation for our empirical investigation, ensuring that our findings are grounded in sound theoretical principles. For ease of reference, we summarize the related notation in Table 1 as follows.

Table 1: Notation used throughout the game description.

| order quantity | demand | cost | price | profit |
|---|---|---|---|---|
| $Q$ | $Y$ | $c$ | $p$ | $\Pi$ |
| shipping cost | clearance price | demand sample | sample size | $\#\{\text{simulations}\}$ |
| $c_S$ | $p_L$ | $\boldsymbol{D}_n = \{Y_i,\, i = 1\ldots, n\}$ | $n$ | $M$ |

**Task 2 (12 points)**

Recall the profit function $\Pi\,(Q, Y; c, p) = p\min\{Q, Y\} - cQ$ as discussed in the first lecture. In our business, we sell goods at a fixed price of $p$ per unit. With rising environmental concerns, more companies are committing to waste reduction. One notable example is the mobile application Too Good To Go, which aims to minimize food waste.

In some cases, when we have surplus or unsold goods, we may consider selling them at a lower price per unit, denoted as $p_L \geq 0$. However, it is important to note that this practice may incur certain costs, such as shipping expenses ($c_S \geq 0$), which are typically borne by the seller. From a purely business perspective, even disregarding the environmental impact, selling unsold perishable products at zero price is preferable to incurring additional fees for waste disposal. This allows us to avoid unnecessary costs and make more efficient use of the resources at our disposal.

(a) Provide a written explanation justifying why $p_L$ should be less than $c + c_S$ in reality.

(b) Revise the profit function to incorporate $p_L$ and $c_S$. This can be represented as $\Pi\,(Q, Y; \tilde{c}, \tilde{p}) = \tilde{p}\min\{Q, Y\} - \tilde{c}Q$, where $\tilde{c}$ and $\tilde{p}$ are functions that depend on $(c, p, c_S, p_L)$. In other words, $(c, p, c_S, p_L)$ affects $\Pi\,(Q, Y; \tilde{c}, \tilde{p})$ through $\tilde{c} = \tilde{c}(c, p, c_S, p_L)$ and $\tilde{p} = \tilde{p}(c, p, c_S, p_L)$.

(c) What is the optimal order quantity $Q^*(F_Y; \tilde{c}, \tilde{p})$ given a CDF $F_Y$, along with $\tilde{c}$ and $\tilde{p}$? It is worth noting that there is no need to provide an extensive mathematical derivation for this, as the solution can be obtained straightforwardly.

(d) Choose a specific distribution and keep the values of $(c, p, p_L)$ fixed. Plot sensitivity graphs (see, e.g., p. 15 of the first lecture) of $Q^*(F_Y; \tilde{c}, \tilde{p})$ and the optimal expected profit with respect to $c_S$. Describe your observations and provide an explanation for the observed patterns.

**Task 3*** (20 points)**

Consider the nonparametric estimator $\widehat{Q}_n^{NP}(\tau) = Y_{(\lceil \tau n \rceil)}$ for some target service level $\tau \in (0, 1)$. This estimator possesses both strengths and weaknesses. On one hand, it offers the advantage of not requiring any distributional assumptions about the demand. On the other hand, Levi et al. (2015, Section 6) have highlighted its potential inaccuracies, particularly when the target service level is high, such as $\tau = 0.99$, and the sample size $n$ is small. To what extent do you agree or disagree with the findings of Levi et al. (2015)? Please provide a detailed explanation of your reasoning and support your opinions by conducting a Monte Carlo simulation study.

To provide some inspiration, a suggested simulation design is outlined on p. 7 below. Before delving into the details of the proposed design, let's first simplify the notation. For simplicity, we can assume $c_S = p_L = 0$. Additionally, for any given target service level $\tau \in (0,1)$, take $(c,p) = (1-\tau, 1)$. Before proceeding, we provide some explanations and introduce the notation as follows.

- Note that the target service level now becomes $(p-c)/p = \tau$. As such, we have $Q^*(\tau) := Q^*(F_Y; c, p) = F_Y^{\leftarrow}(\tau)$, which is the theoretically optimal quantity for $\tau$.

- Assume $F_Y(\cdot) = G_{\boldsymbol{\theta}}(\cdot) = G(\cdot \,|\, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is an unknown parameter. For any $\tau \in (0,1)$, we define $\widehat{Q}_n^P(\tau) := \widehat{Q}_n^P(G_{\widehat{\boldsymbol{\theta}}_n}; \tau) = G^{\leftarrow}\big(\tau \,\big|\, \widehat{\boldsymbol{\theta}}_n\big)$, where $\widehat{\boldsymbol{\theta}}_n$ denotes the maximum likelihood estimator of $\boldsymbol{\theta}$. This benchmark estimator, $\widehat{Q}_n^P(\tau)$, will serve as our reference for comparison purposes.

- Consider the estimate $\widehat{Q}_n^{k,j}(\tau)$, where $k$ represents the estimation method (either $NP$ or $P$), and $j$ denotes the $j_{th}$ estimated value in the $M$ Monte Carlo repetitions. For example, $\widehat{Q}_n^{P,10}(\tau)$ corresponds to the parametric estimate of the order quantity in the $10_{th}$ Monte Carlo iteration.

- Let $R(Q;\tau) := \mathbb{E}_Y\left[\Pi(Q, Y; c = 1-\tau, p = 1)\right]$ be the profit function, where the argument $Q$ is a function that depends on $\tau \in (0,1)$. Namely, $Q = Q(\tau)$, as explained on p. 17 of the first lecture. For instance, $R(Q^*;\tau)$ is the maximum expected profit. Similarly, for any given $\tau \in (0,1)$, $R\big(\widehat{Q}_n^{NP};\tau\big)$ represents the resulting profit using $\widehat{Q}_n^{NP}(\tau)$. Note that, for any $\tau \in (0,1)$, $R(Q^*;\tau)$ is a constant, while $R\big(\widehat{Q}_n^{NP};\tau\big)$ is random because $\widehat{Q}_n^{NP}(\tau)$ is random.

To evaluate and compare the performance of $\widehat{Q}_n^{NP}(\tau)$ and $\widehat{Q}_n^P(\tau)$, we suggest three criteria:

- Empirical root mean squared error (RMSE):

$$\mathrm{RMSE}_n^k(\tau) = \left[M^{-1}\sum_{j=1}^{M}\left(\widehat{Q}_n^{k,j}(\tau) - Q^*(\tau)\right)^2\right]^{1/2}, \qquad k \in \{NP, P\}. \tag{1}$$

  Of particular interest is the ratio:

$$\mathrm{RMSE}_n^{NP}(\tau) \big/ \mathrm{RMSE}_n^P(\tau), \tag{2}$$

  where the denominator is considered as the benchmark. If the ratio is greater than 1, it indicates that $\widehat{Q}_n^P(\tau)$ is more accurate than $\widehat{Q}_n^{NP}(\tau)$ for some $\tau \in (0,1)$.

- Empirical service level (SL):

$$\mathrm{SL}_n^k(\tau) = M^{-1}\sum_{j=1}^{M} \mathbb{1}\left\{\widehat{Q}_n^{k,j}(\tau) \geq y_n^j\right\}, \qquad k \in \{NP, P\}, \tag{3}$$

  where $\mathbb{1}\{\cdot\}$ is an indicator function that takes the value 1 if the condition inside the brackets is fulfilled, and 0 otherwise. Moreover, $y_n^j$ represents the observed value of $Y_n \in \boldsymbol{D}_n$ in the $j_{th}$ Monte Carlo repetitions. Note that the random demand sample $\boldsymbol{D}_n$ (defined in Table 1) comprises $n$ i.i.d. random variables $Y_1, \ldots, Y_n$; we simply use the observation for the last random variable. For the given values of $(c, p, c_S, p_L) = (1-\tau, 1, 0, 0)$, the empirical service level should match the target service level $\tau$.

- Empirical profit loss ratio (PLR):

$$\mathrm{PLR}_n^k(\tau) = M^{-1}\sum_{j=1}^{M} \left|\frac{R(Q^*;\tau) - R\big(\widehat{Q}_n^{k,j};\tau\big)}{R(Q^*;\tau)}\right|, \qquad k \in \{NP, P\}. \tag{4}$$

  Similarly, one can compare the performance of the estimators by considering the ratio:

$$\mathrm{PLR}_n^{NP}(\tau) \big/ \mathrm{PLR}_n^{P}(\tau). \tag{5}$$

  If the ratio is greater than 1, then $\widehat{Q}_n^{NP}(\tau)$ has higher profit loss compared to $\widehat{Q}_n^{P}(\tau)$.

In response to the findings by Levi et al. (2015), we can conduct a simulation study to evaluate the performance of the estimators by varying $\tau$ and $n$. Here is a suggested simulation design:

(a) To compare the performance of $\widehat{Q}_n^{NP}(\tau)$ and $\widehat{Q}_n^{P}(\tau)$ under various scenarios, we can consider different demand distributions. For instance, one may assume the true demand follows: (i) $\mathcal{N}(\mu, \sigma^2)$ with $\mu = 115$, $\sigma = 10$, and/or (ii) $\mathcal{LN}(\mu, \sigma^2)$ with $\mu = 6$ and $\sigma = 0.6$.

(b) Utilize the criteria (2), (3), and (5) to evaluate the performance. Consider different sample sizes, $n \in \{10, 50, 100, 200\}$, and target service levels, $\tau \in \{0.01, 0.05, 0.1, 0.3, 0.5, 0.7, 0.9, 0.95, 0.99\}$.

(c) Furthermore, to expand the analysis, one can explore different parameter values for the parametric distribution of demand. For example, consider the $\mathcal{N}(\mu, \sigma^2)$ distribution with varying parameter values, such as $(\mu, \sigma) \in \{(110, 10), (115, 15)\}$.

These suggestions are intended to inspire your simulation design. For a detailed example of creating, discussing, and analyzing a simulation design, you may refer to Beutner et al. (2023, Section 6) or Friedrich and Lin (2024, Section 5). If you choose to utilize any of the proposed criteria, namely (2), (3), and (5), it is important to explain why they are meaningful in words.

We welcome any additional comparison criteria and constructive simulation designs that you may propose. It is important to note that the numerical evaluation of $R(Q;\tau)$, whether through simulation or using the built-in function "quad", can be computationally intensive. To reduce computational cost, taking $M = 500$ or $M = 1000$ repetitions is usually sufficient for this task.

**Task 4** (12 points)**

In Task 3, we mentioned that the nonparametric estimator of order quantity has both strengths and weaknesses compared to the parametric estimator. If your simulation study yields favorable results in that exercise, you should be able to observe some advantages of the parametric estimator. However, it is essential to recognize the weaknesses of the parametric estimator as well. As discussed in the lectures, the parametric estimator can be sensitive to the distribution specification. Levi et al. (2015, Section 6) have commented that

> "if the true demand distribution is nonnormal, then fitting the sample to a normal distribution might result in a suboptimal order quantity."

To what extent do you agree or disagree with the comment of Levi et al. (2015)? Based on your simulation study for Task 3, conduct additional simulations to support your own opinions. You can begin by generating the demand using a non-normal distribution and subsequently fitting it with a normal distribution. You may also consider generating the data using a normal distribution and then fitting it with a non-normal distribution.

## 2.3 Empirical tasks

The theoretical insights obtained previously contribute to addressing real-life inquiries. In this task category, we examine a specific subset of a large dataset obtained from a local bakery chain store in Vilnius, Lithuania.

**Background information**

This subset of the dataset comprises daily demand data for a traditional Lithuanian dessert from four different stores, covering the period from *11-05-2016* to *29-05-2024*. The dataset includes the following information:

- The "weekday" column specifies the day of the week, with Monday=1, Sunday=7.
- The subsequent columns provide the daily demand for the traditional dessert, measured in 0.25 kg, for each of the four stores.

Two of these stores are situated on main streets in Vilnius, while the other two are near central bus/train stations. Each location has its own opening date, as shown in the table below.

| location | main street A | main street B | station A | station B |
|---|---|---|---|---|
| Opening date | 18-04-2016 | 30-01-2023 | 20-08-2018 | 06-07-2022 |

Table 2: Cost and price information: All values are expressed in euros per 0.25 kg.

| location | cost ($c$) | price ($p$) | shipping cost ($c_S$) | clearance price ($p_L$) |
|---|---|---|---|---|
| main street A | 3.85 | 4.64 | 0.11 | 0.15 |
| main street B | 3.42 | 4.64 | 0.08 | 0.15 |
| station A | 4.16 | 4.64 | 0.08 | 0.15 |
| station B | 3.32 | 4.64 | 0.09 | 0.15 |

The first store, labeled as main street A, opened on 18-04-2016 and is located on a vibrant main street in the heart of the old town. This street is renowned for its numerous tourist attractions, cultural landmarks, and fashionable boutiques. On Friday afternoons and Saturdays, it attracts a significant number of visitors. However, it is important to note that data collection for this store began on 11-05-2016 due to technical issues during the initial period of its opening. The second store, labeled as main street B, opened on 30-01-2023 and is situated on a prominent main street characterized by the presence of large companies. It benefits from the constant flow of people commuting to work in the area. Additionally, there is a store labeled as station A, which is located close to the central train station. This location attracts a high volume of people passing through for their daily commute. On the other hand, the store labeled as station B is situated near the old town, offering convenient access to tourists and locals exploring the historic area.

Moreover, the stores located on main street A and station A experienced a significant impact from the COVID-19 pandemic, starting around March 2021. The demand for their products declined sharply during this period. However, by approximately March 2022, the demand began to recover, gradually returning to its original level. The remaining two stores were opened after the COVID-19 pandemic, and therefore, they have experienced relatively stable demand, unaffected by the pandemic-related shocks experienced by the other stores.

Given the changing economic environment, including factors such as inflation, it is not surprising that costs and thus prices have undergone multiple small increases. It is important to note that cost values encompass a range of elements, such as materials, rents, electricity, staff salaries, and more, and thus are subject to variation over time. Nevertheless, to simplify your analysis, we will use the average cost as a representative measure for the past years. Furthermore, as mentioned earlier, the chain store has established a close collaboration with a local feed mill. On a daily basis, the bakery store collects surplus products and arranges for their own delivery to the feed mill. The transportation fee ($c_S$) may vary depending on the location of each shop, but the clearance price ($p_L$) offered by the feed mill remains identical. For convenience, we summarize the information regarding cost and price in Table 2 above.

**Task 5 (12 points)**

Import and describe the data (BakeryData2024_Vilnius.xlsx). This may include:

(a) Calculate descriptive statistics and create box plots to examine the daily demand for each store. One may look at measures like sample mean, sample variance, and the numbers of observations.

(b) Plot histograms to visualize the distribution of daily demand for each store.

(c) Create time series plots that display the daily demand on the $y$-axis and the corresponding dates on the $x$-axis.

When reviewing the plots, do they align with the background information? Are there any discernible patterns in the demand data, especially during weekends?

**Task 6* (24 points)**

Based on the provided background information, estimate the optimal quantities for the short run (near future) using both parametric and nonparametric approaches. For instance, we consider the out-of-sample dates 31-05-2024 (Friday), 01-06-2024 (Saturday), and 02-06-2024 (Sunday) as the short term.

- *Parametric framework:*

(a) Select at least two stores for your analysis, ensuring that the main street A store is included.

(b) Explain your approach to specifying/choosing the demand distributions, based on your results for Task 5.

(c) Provide point estimates and intervals for the theoretically optimal quantities.

- *Nonparametric framework:*

(d) Calculate point estimates and obtain intervals for the theoretically optimal quantities of the same stores selected above.

Compare thoroughly the similarities and differences between the parametric and nonparametric methods. You may link your discussions to the findings from Tasks 3 and 4. It is important to note that, given $c_S \neq 0$ and $p_L \neq 0$, the profit function $\Pi(Q, Y; \tilde{c}, \tilde{p})$ derived in Task 2 should be employed.

### Task 7** (14 points)

Due to various factors, including the Russia-Ukraine War, costs have experienced a significant surge in recent periods. In light of this, the chain store anticipates a substantial, approximate 25% increase in costs over the next 6 months, which is usually considered as the long run compared to Task 6. To counterbalance these escalating expenses, the store's managers have decided to strategically raise the selling price ($p$). Their market research has revealed a crucial insight: $r_Y$, which represents the ratio by which the expectation of demand drops, is influenced by $r_p$, the ratio of price increase. This relationship can be approximately captured by the following equation:

$$r_Y = \left\{ \left[ 1 + \exp\left( 6 - \frac{r_p}{10} \right) \right]^{-1} - 0.0025 \right\} \times 100. \tag{6}$$

For example, if the price increases by 50%, the expected demand decreases by approximately 26.6441%, calculated as follows:

$$r_Y = \left\{ \left[ 1 + \exp\left( 6 - \frac{50}{10} \right) \right]^{-1} - 0.0025 \right\} \times 100 \approx 26.6441.$$

Offer some recommendations to the managers regarding the extent to which the price should be increased in the next 6 months to maintain a similar profit level as the present, assuming other factors remain unchanged. Do you have any other recommendations or suggestions that the managers should take into consideration?

## 2.4 Critical review

### Task 8 (6 points)

There are some restrictions associated with the methods introduced in the course. Admittedly, given your current knowledge, the analysis in Sections 2.2 and 2.3 may also have multiple limitations. Write a concise critical review of your analysis within a maximum of one A4 page. You may pay attention to the following aspects:

(a) Discuss the limitations of the methods if they are not fully appropriate for the data.

(b) Elaborate on the assumptions you have made, explaining why you consider them necessary and whether they impose any restrictions.

(c) Although life must be lived forwards, it "can only be understood backwards" as the existentialist philosopher Søren Kierkegaard said. Looking back, you may have gained new insights into NVPs and related methods. For instance, in your writing assignment, you were asked to compare Monte Carlo simulation and bootstrap methods. Has your understanding improved as a result of this course?

(d) What recommendations can you offer for future research? Are there any potential areas for further exploration and extension?

These suggestions are merely provided as guidelines for your focus in this task. You are welcome to explore other discussions as well. What matters most is that you thoroughly explain your reasons. A structure that reflects this approach could be exemplified by phrases such as "We find that ... because...".

# 3 Bonus point task***** (50 points)

In our course, we have assumed that the demand is i.i.d., implying that today's demand is independent of yesterday's demand. However, this assumption may not hold in real life. If the demand is not independent over time, how would you determine the optimal order quantity and construct the corresponding confidence intervals? Share your thoughts and provide numerical evidence to support your ideas. For example, you could consider using Monte Carlo simulations to explore your solutions in different scenarios.

As previously mentioned, there are no strict page limits for the answers to the bonus point task. However, it is highly recommended to keep your response within a maximum of 10 pages. Please note that your answer to this task will only be evaluated if you achieve a score of 80 or higher (out of 100) in the main tasks given in Section 2.

**Enjoy and Success!**

# References

Beutner, E., Y. Lin, and S. Smeekes (2023). GLS estimation and confidence sets for the date of a single break in models with trends. *Econometric Reviews 42*(2), 195–219.

Friedrich, M. and Y. Lin (2024). Sieve bootstrap inference for linear time-varying coefficient models. *Journal of Econometrics 239*(1), 105345.

Levi, R., G. Perakis, and J. Uichanco (2015). The data-driven newsvendor problem: New bounds and insights. *Operations Research 63*, 1294–1306.