



國立臺灣大學工學院工業工程學研究所

碩士論文

Graduate Institute of Industrial Engineering

College of Engineering

National Taiwan University

Master's Thesis

長短期記憶神經網路之消費者偏好學習與

動態訂價運用

Adopting Long Short-Term Memory Neural Network for
Consumer Preference Learning and Dynamic Pricing

王文謙

Wen-Chian Wang

指導教授：吳政鴻 博士

Advisor: Cheng-Hung Wu, Ph.D.

中華民國 113 年 1 月

January 2024

誌謝



時間飛逝，兩年多的碩士生涯迎來了尾聲，也代表十幾年的學生身分要暫時告一段落了。依然記得大四那年埋首書中，和考古題奮鬥的日子依然歷歷在目，如今即將踏入下一個人生階段，我想感謝在研究所中陪伴著我，幫助過我的人們。

首先，我要感謝我的父母，不管在生活上、學業上總是一路支持著我，在我失落、徬徨、找不到方向的時候陪伴著我，給予我不同面向的觀點；需要做出任何決定時和我一起分析利與弊，而不是直接叫我該怎麼做。如果沒有您們，不會有現在的我，謝謝我的家人作為我堅強的後盾，給我無限的支持與愛。

再來我要感謝吳老師，很幸運能作為您的學生，儘管我在碩一時非常的混，咪挺常常用很爛的東西想要打發過去，吳老師也都沒有放棄我，時常叮囑我研究進度需要加緊腳步，儘管最後還是多陪老師半年的時間~這些做研究的日子裡，不僅學習到老師專業的知識，更重要的是教導我們獨立思考、解決問題的能力，以及報告的小細節，謝謝吳老師的關心以及指教。

我要感謝吳 lab 的朋友們以及其他實驗室的朋友們，一起討論艱難的作業和研究，也會分享生活的酸甜苦辣，謝謝各位豐富我研究所的生活。

我要感謝我的女友這些年來的陪伴，在我遇到研究低潮，面試不太順利的時候，都會不斷鼓勵我，並且陪伴我修改我的問題，讓我變成更好的人。

最後我要感謝一路上曾經幫助過我，很幸運能認識大家，祝福各位一帆風順，心想事成。

王文謙 謹致於臺灣大學

2024 年 冬

中文摘要及關鍵詞



需求學習(Demand Learning)可以幫助供應商了解消費者喜好來達到滿足市場需求，但往往受限於歷史資料不足及環境變動快速等問題，無法得到最佳的效果。因應此問題，朱婉琳(2019)提出一種結合動態規劃(Dynamic Programming)與深度學習(Deep Learning)之技術應用在動態訂價(Dynamic Pricing)與需求學習中，該模型使用長短期記憶神經網路(Long Short-Term Memory neuron network, LSTM)對各種市場環境進行消費者需求學習，在面對新環境時透過歷史資料估計新環境下消費者支付意願(Willing to Pay, WTP)之參數，透過動態訂價求解出最佳決策並逼近最佳利潤。

然而，朱婉琳之研究僅透過神經網路學習消費者支付意願分布之平均值，並假設支付意願服從標準差為固定值之常態分布(Normal Distribution)，並不符合實際市場中的狀況:消費者支付意願為服從任意常態分布，即常態分布之平均值及標準差值皆是變動的，亦或是支付意願服從其他各種不同分布。因此，本研究將使用神經網路模型學習消費者支付意願分布，其中透過KL Divergence作為神經網路的損失函數(Loss)去學習常態分布之平均值及標準差，或是學習伽瑪分布的參數，更符合實際市場銷售情形，並透過動態訂價求解，以達到供應商販售時效性商品之最大化利潤。

最後透過模擬訂價驗證在各式各樣的市場環境下，本研究提出的模型僅需要用很少的成本學習消費者的需求，且模擬訂價的結果與動態規劃求解最佳利潤的差異很少，結果證實，本研究之模型能在未知的環境下擁有良好的能力。

關鍵字:動態規劃、動態訂價、需求學習、深度學習、長短期記憶神經網路

英文摘要及關鍵詞

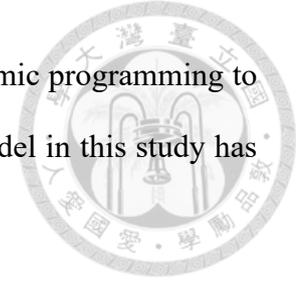


Demand Learning can help suppliers understand consumer preferences to meet market demand. However, it is often limited by insufficient historical data and rapid environmental changes, and cannot achieve the best results. In response to this problem, Zhu(2019) proposed a technology that combines dynamic programming and deep learning to be used in dynamic pricing and demand learning. This model uses Long Short-Term Memory neuron network(LSTM) conducts consumer demand learning in various market environments, and uses historical data to predict consumer willingness to pay (WTP) in the new environment through historical data. Dynamic pricing solves the best decision and approaches the best profit.

However, Zhu's research only used neural networks to learn the average value of consumers' WTP distribution, and assumed that WTP obeys a normal distribution with a fixed standard deviation, which is not consistent with the situation in the actual market: consumers The WTP is subject to any normal distribution, that is, the mean and standard deviation of the normal distribution are changing, or the WTP is subject to various other different distributions. Therefore, this study will use a neural network model to learn the distribution of consumers' willingness to pay, using KL Divergence as the loss function of the neural network to learn the mean and standard deviation of the normal distribution, or learn the gamma distribution The parameters are more in line with the actual market sales situation and are solved through dynamic pricing to maximize the supplier's profit from selling time-sensitive goods.

Finally, it is verified through simulated pricing that in various market environments, the model proposed in this study only needs to learn consumer needs at a very small cost,

and the results of simulated pricing are different from those of dynamic programming to solve the optimal profit. Very rarely, the results confirm that the model in this study has good capabilities in unknown environments.



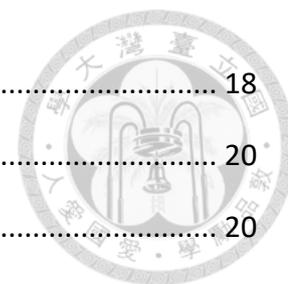
Keywords: Dynamic programming 、 Dynamic pricing 、 Demand learning 、 Deep learning 、 LSTM neural network

目次



誌謝	i
中文摘要及關鍵詞	ii
英文摘要及關鍵詞	iii
目次	v
圖次	vii
表次	x
第一章 緒論	1
1.1 研究背景與動機	1
1.2 研究目的	2
1.3 研究方法與流程	2
第二章 文獻回顧	4
2.1 動態訂價	4
2.2 人工智慧在動態訂價之應用	8
2.2.1 機器學習在動態訂價之應用	8
2.2.2 神經網路在動態訂價之應用	9
2.2.3 時間序列預測之神經網路	9
2.2.4 損失函數的選擇	11
2.3 小結	12
第三章 長短期記憶神經網路之架構	14
3.1 長短期記憶神經網路之介紹	14
3.2 消費者購買資料蒐集	16
3.2.1 消費者購買資料模擬	16

3.2.2 資料特徵萃取	18
3.2.3 模擬資料劃分	20
3.3 神經網路架構	20
第四章 動態訂價模型	25
4.1 模型基本架構與假設	25
4.2 動態訂價模型	26
第五章 模型測試與模擬驗證	30
5.1 WTP 服從常態分布	30
5.1.1 參數設置	30
5.1.2 長短期記憶神經網路模型測試	31
5.1.3 模擬驗證流程	33
5.1.4 模擬驗證結果	36
5.2 WTP 服從伽瑪分布	47
5.2.1 參數設置	47
5.2.2 長短期記憶神經網路模型測試	48
5.2.3 模擬驗證流程	50
5.2.4 模擬驗證結果	51
第六章 結論與未來方向	66
6.1 結論	66
6.2 未來研究方向	66
參考文獻	68



圖次



圖 1 研究流程圖	3
圖 2 LSTM 神經元內部結構圖	14
圖 3 消費者購買資料模擬流程圖	18
圖 4 單筆 LSTM 輸入資料型態	19
圖 5 常態市場平均值及標準差落在每一種數值的機率	20
圖 6 Target Replication Strategy	21
圖 7 LSTM 神經網路架構圖(常態分布)	22
圖 8 購買機率計算示意圖	27
圖 9 估計 250 期銷售時程之常態分布平均值	32
圖 10 估計 250 期銷售時程之常態分布標準差值	32
圖 11 LSTM 模型之訓練常態分布 Loss 圖	37
圖 12 模擬標準差為 10 之估計 250 期銷售時程的 μ 和 σ	37
圖 13 模擬標準差為 10 之第一次模擬之消費者願付價格機率密度函數柱 狀圖	38
圖 14 模擬標準差為 10 之第一次模擬 71 種平均之模擬利潤	39
圖 15 模擬標準差為 10 之平均為 70 的 30 次模擬利潤	39
圖 16 模擬標準差為 10 之偏好學習成本圖	40
圖 17 模擬標準差為 20 之估計 250 期銷售時程的 μ 和 σ	41
圖 18 模擬標準差為 20 之第一次模擬之消費者願付價格機率密度函數柱 狀圖	41
圖 19 模擬標準差為 20 之第一次模擬 71 種平均之模擬利潤	42
圖 20 模擬標準差為 20 之平均為 70 的 30 次模擬利潤	42



圖 21 模擬標準差為 20 之偏好學習成本圖	43
圖 22 模擬標準差為 30 之估計 250 期銷售時程的 μ 和 σ	43
圖 23 模擬標準差為 30 之第一次模擬之消費者願付價格機率密度函數柱 狀圖	44
圖 24 模擬標準差為 30 之第一次模擬 71 種平均之模擬利潤.....	45
圖 25 模擬標準差為 30 之平均為 70 的 30 次模擬利潤.....	45
圖 26 模擬標準差為 30 之偏好學習成本圖	46
圖 27 估計 250 期銷售時程之伽瑪分布形狀參數值	48
圖 28 估計 250 期銷售時程之伽瑪分布尺度參數值	49
圖 29 估計 250 期銷售時程之伽瑪分布形狀參數乘上尺度參數之值	49
圖 30 LSTM 模型之訓練伽瑪分布 Loss 圖	51
圖 31 模擬尺度參數為 2 之估計 250 期銷售時程的 α 和 θ	52
圖 32 模擬尺度參數為 2 之估計 250 期銷售時程的 $\alpha*\theta$	52
圖 33 模擬尺度參數為 2 之第一次模擬之消費者願付價格機率密度函數 柱狀圖	53
圖 34 模擬尺度參數為 2 之第一次模擬 21 種形狀參數之模擬利潤.....	54
圖 35 模擬尺度參數為 2 之形狀參數為 10 的 30 次模擬利潤.....	54
圖 36 模擬尺度參數為 2 之偏好學習成本圖	55
圖 37 模擬尺度參數為 3 之估計 250 期銷售時程的 α 和 θ	55
圖 38 模擬尺度參數為 3 之估計 250 期銷售時程的 $\alpha*\theta$	56
圖 39 模擬尺度參數為 3 之第一次模擬之消費者願付價格機率密度函數 柱狀圖	56
圖 40 模擬尺度參數為 3 之第一次模擬 21 種形狀參數之模擬利潤.....	57
圖 41 模擬尺度參數為 3 之形狀參數為 10 的 30 次模擬利潤.....	57
圖 42 模擬尺度參數為 3 之偏好學習成本圖	58

圖 43 模擬尺度參數為 4 之估計 250 期銷售時程的 α 和 θ	59
圖 44 模擬尺度參數為 4 之估計 250 期銷售時程的 $\alpha*\theta$	59
圖 45 模擬尺度參數為 4 之第一次模擬之消費者願付價格機率密度函數 柱狀圖	60
圖 46 模擬尺度參數為 4 之第一次模擬 21 種形狀參數之模擬利潤.....	61
圖 47 模擬尺度參數為 4 之形狀參數為 10 的 30 次模擬利潤.....	61
圖 48 模擬尺度參數為 4 之偏好學習成本圖	62
圖 49 模擬尺度參數為 5 之估計 250 期銷售時程的 α 和 θ	62
圖 50 模擬尺度參數為 5 之估計 250 期銷售時程的 $\alpha*\theta$	63
圖 51 模擬尺度參數為 5 之第一次模擬之消費者願付價格機率密度函數 柱狀圖	63
圖 52 模擬尺度參數為 5 之第一次模擬 21 種形狀參數之模擬利潤.....	64
圖 53 模擬尺度參數為 5 之形狀參數為 10 的 30 次模擬利潤.....	64
圖 54 模擬尺度參數為 5 之偏好學習成本圖	65

表次



表 1 數學符號定義表	18
表 2 WTP 為常態分布之參數設置	31
表 3 KL Divergence 值比較表	33
表 4 模擬驗證參數表	35
表 5 WTP 為伽瑪分布之參數設置	47

第一章 緒論



動態訂價是指供應商根據市場需求和自身供應能力，適時調整商品或服務之最佳價格，以實現收益最大化的訂價策略。然而，隨著時代的進步，市場上的需求變化更為快速，特別是對於販售時效性商品之供應商而言，無法即時蒐集到銷售資料，市場環境中又充滿不確定性，導致無法做出優良的初始決策。

本研究將提出結合動態規劃與需求估計技術應用於動態訂價中，以解決需求學習中歷史資料不足以及環境不確定性之問題，並同時又能與最佳利潤相去不遠。

本章節首先描述研究背景與動機，再來介紹本研究目的，最後則說明研究中使用之方法及研究流程。

1.1 研究背景與動機

需求學習能夠幫助販售時效性商品的供應商在未知的市場環境中進行學習，並幫助供應商做出良好的訂價決策。然而隨著時代與網路技術的進步，消費者的喜好與需求變化快速，使供應商在尚未蒐集到足夠多的銷售資料前銷售時程就已經結束，造成需求學習的歷史資料不足；在缺乏歷史資料情況下，供應商往往會使用類似產品做為決策參考，但是不同產品會面臨不同環境及需求，且市場環境會受到許多未知因素影響，假使沒有在銷售時程開始前做出準確的訂價策略，並隨著銷售時程開始做出迅速及準確的反應處理環境不確定性，會造成供應商利潤重大損失。

動態規劃是透過貝爾曼方程式(Bellman Equation)及馬可夫決策過程(Markov decision process, MDP)求解出在不同狀況下的最佳策略(Bellman, 1954)，求解過程中需要確定系統中每個參數才能計算出轉移機率及回饋函數。然而，歷史資料不足以及市場環境不確定性往往導致參數未知或是充滿不確定性，並且當任何參數



發生變化時，將會導致轉移機率及回饋函數發生改變，造成最佳解的改變。因此當研究中有任何參數改變，都必須重新建構一次動態規劃模型，再進行求解，不僅費時也費力。

1.2 研究目的

為解決上述困難，朱婉琳(2019)提出一種能夠處理環境不確定性且能做出良好初始決策的需求學習方法，以動態估計參數的方式解決動態規劃模型假設問題。然而，朱婉琳之研究僅透過神經網路學習消費者支付意願分布的平均值，並假設支付意願服從常態分布且標準差為固定值，並不符合實際市場中的狀況:消費者支付意願服從任意常態分布，即常態分布之平均值及標準差皆是變動的，亦或是支付意願服從其他各種不同分布。因此，本研究將使用神經網路模型學習消費者支付意願分布，其中透過 KL Divergence 作為神經網路的損失函數(Loss)去學習常態分布之平均值及標準差，或是學習伽瑪分布的參數，更符合實際市場銷售情形，並透過動態訂價求解，最後使用模擬驗證本研究模型之效力，使用很少的成本學習消費者需求且達到供應商販售時效性商品之最大化利潤。

本研究將結合深度學習及動態規劃，訓練出能夠辨識新環境且做出良好初始決策，並即時進行動態決策優化的模型，該模型著重於環境識別及即時需求學習上，以販賣一種時效性商品之市場作為本研究之環境。

1.3 研究方法與流程

本小節將介紹本研究所使用之方法與流程。本研究結合需求估計與長短期記憶神經網路識別新環境下的系統參數，並將系統參數帶入動態規劃中求解最佳訂價，並比較真實利潤，計算學習真實利潤所需要花費的成本。

本研究首先模擬大量消費者購買資料，從中萃取特徵作為神經網路的訓練資料和後續驗證的測試資料，訓練出可識別未知環境的 LSTM 模型，並在模擬驗證



時將估計之市場環境參數帶入動態規劃求解，得到模擬利潤，最後將模擬利潤與使用已知環境參數得到之真實利潤做比較，計算偏好學習成本，證明本研究方法之可行性，流程圖如圖 1 所示：

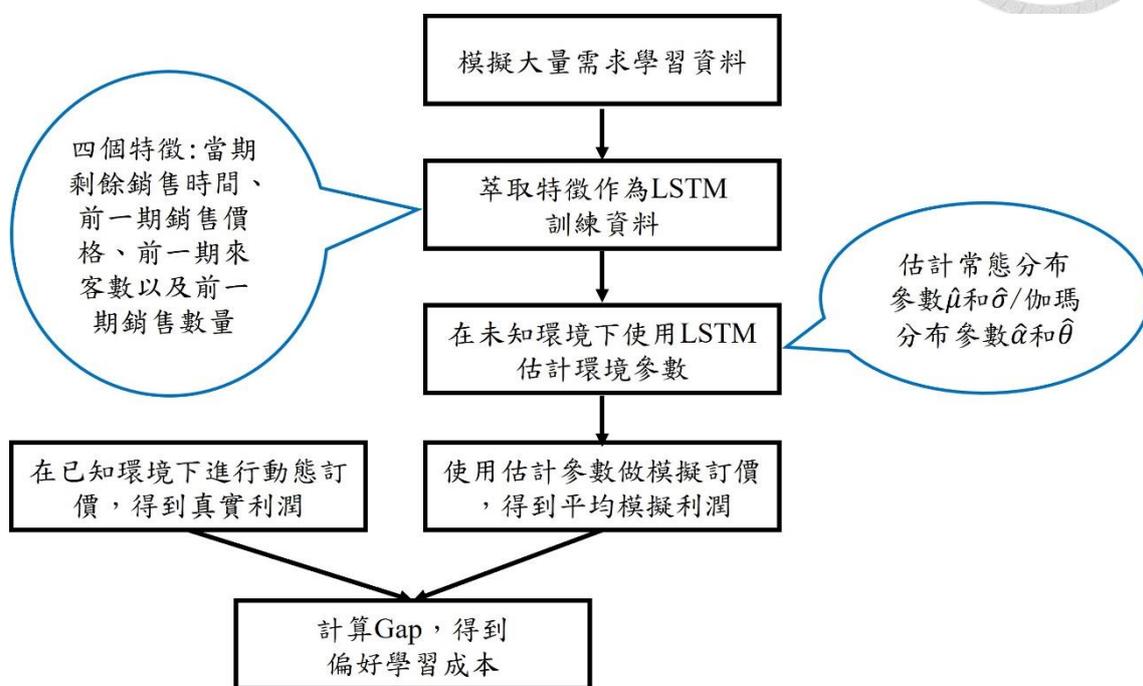


圖 1 研究流程圖

第二章將進行需求學習、動態規劃與動態訂價和神經網路及其應用之文獻回顧。第三章介紹長短期記憶神經網路之架構以及訓練資料的模擬。第四章介紹動態規劃模型的假設及建構。第五章說明參數設定以及模擬驗證。第六章進行結論並探討未來研究方向。

第二章 文獻回顧



本章節會回顧及探討現有的相關文獻，首先會先探討動態訂價之相關文獻；第二節探討機器學習及神經網路在動態訂價之應用，並探討時間序列預測之神經網路與損失函數的選擇；最後為文獻回顧的總結。

2.1 動態訂價

本節將介紹動態訂價的發展背景與研究歷史，並依照研究主題做分類，並說明現有文獻中面臨到的問題與挑戰。

動態規劃在 1950 年代早期由數學家 Richard Bellman 提出(Bellman, 1954)，可用於解決不確定情況下的各種順序決策問題，在每一種狀態中都進行最佳決策並達到最優解。動態規劃在許多領域上都能應用，如經濟、作業研究、管理科學、工程相關以及人工智慧上，是解決複雜決策問題的強力工具(Rust&John., 2019)。動態規劃應用於訂價方面問題稱之為動態訂價，動態訂價可以輕鬆且頻繁地調整產品或服務的價格，使其達到最佳銷售價格。

動態訂價之研究主題龐大，與需求學習相關是一大研究重點。需求學習是供應商在現今市場中制訂訂價策略的有效方法，透過歷史銷售狀況和即時銷售數據加深對消費者需求的了解，即時優化訂價策略。早期文獻中皆假設消費者需求為已知，然而這和真實情況並不相符，因此後續開始研究需求為未知之動態訂價模型，並且使用消費者歷史購買資料對於需求函數進行估計。而與需求學習相關之問題依照庫存限制又可以分成有限庫存以及無限庫存；另外也有動態訂價考量消費者行為，以及考量競爭對手作為訂價之影響。

➤ 已知需求函數



(Cournot, 1927)被認為是第一位使用數學函數表達產品價格與消費者需求之關係，使用確定性需求函數解決靜態獨佔訂價問題。然而這些確定性假設與實際情況並不相符，需求函數應該為隨機變數而不是確定性的。(Mills, 1959)最先提出此想法，研究在有限時程內銷售多種產品的供應商如何設置最佳化商品生產數量以及銷售價格，並假設需求是平均為零的隨機項與確定性價格函數的總和。(Nevins, 1966)考量一個動態理論，商品數量與商品售價是每個時期的決策變數，並透過大量數值計算得到近似最佳解。(Sandmo, 1971)研究在需求不確性情況下，供應商對風險偏好如何影響最佳訂價決策。

儘管上述研究將需求建模為隨機變數，然而其預期需求(商品售價之函數)仍為供應商已知，這是不切實際的，因此後續的文獻開始探討商品售價與預期需求之間不確定性的訂價策略，其中需求學習依照庫存數量可以分為無限庫存以及有限庫存之訂價問題。

➤ 未知需求函數與無限庫存

(Aoki, 1973)使用隨機自適應控制理論的方法(Stochastic Adaptive Control Theory)，需求函數取決未知參數，並透過動態訂價計算最優近似策略。(Wruck, 1989)考慮了兩階段模型中易腐爛產品和非易腐爛產品的最優訂價，透過貝氏學習方式學習未知參數。(Qu, 2013)假設需求分布服從伯努利分布(Bernoulli Distribution)且期望價格為 Logit 函數，考慮一種近似貝葉斯統計模型估計售價的方法，並將需求不確定性納入決策中。(Venezia, 1978)研究一種具有未知需求參數的線性模型，並假設其中一個參數之行為類似於隨機遊走(Random Walk)。然而這些研究都存在明顯缺點，即最優策略計算之困難性與複雜性。

(Carvalho, 2005)考慮在需求函數未知下，透過觀察歷史價格估計出消費者需求，隨著銷售時程更新售價達到利潤最大化。(Morales-Enciso et al., 2012)在無記憶市場環境下，探討消費者對於不同價格的反應，估計出消費者支付意願並使用啟發式訂價策略最大化整體利潤。(Eren & Maglaras, 2010)探討在穩健環境下研究



動態訂價，並證明在有限銷售時程內銷售無限庫存之商品，最好的訂價策略為撇脂定價法(Price-Skimming Strategy)

➤ 未知需求函數與有限庫存

(Kincaid & Darling, 1962)可能是第一個研究在有限時程內銷售有限庫存對於最優訂價策略影響的人。(Aviv et al., 2005)探討在需求不確定下，供應商可以透過銷售數據解決其部分不確定性，且使用主動學習啟發式訂價策略使利潤最大化。(Farias, 2010)研究庫存有限且需求不確定之訂價問題，提出一種衰減平衡(Decay Balancing)的訂價啟發式演算法，並表明與貪婪啟發式演算法相比可帶來顯著的利潤。(Xiong, 2010)不依賴歷史資料解決需求不確定性，而是使用模糊理論(Fuzzy Theory)對需求不確定性進行建模，並使用遺傳演算法求解問題。(Avramidis, 2011)研究發現使用消費者到達率以及產品銷售數量可以精準估計需求函數。(Chen, 2014)透過雙邊審查觀察(Two-Sided Censored Observations)方法估計消費者支付意願分布，並使用三種啟發法得到近似最優解。(den Boer & Bert Zwart, 2015)研究需求分布確定但參數不確定性的情況下，產品在有限時程內銷售有限庫存的即時學習策略與動態訂價問題，並使用最大概似估計(Maximum-Likelihood Estimation)對未知參數進行估計。然而這些啟發式算法與數學估計皆有一大通病，即計算最優策略非常困難與複雜，且無法獲得精準的需求估計。

(Gallego et al., 1994)假設需求服從卜瓦松分布(Poisson Distribution)且消費者到達率取決於公布的銷售價格情況下，探討有限時程及庫存內最優訂價策略。(Chen et al., 1999)研究消費者支付意願分布未知下，供應商透過銷售數據推測其需求分布。(Lin, 2006)研究消費者支付意願分布未知下，使用即時銷售數據來調整初始估計的消費者到達率來更精確地估計未來的需求分布，以最大限度地提高預期的總收入。(Escobari, 2014)使用機票銷售價格與銷售數量估計消費者需求。也有文獻使用 Logit 模型估計消費者支付意願分布(Hole et al., 2012)，(Wen et al., 2017)

使用連續多項 Logit 模型估計消費者需求。然而其中 Logit 模型有一大缺點，即 Logit 模型估計過程會使隨機變量參數更為複雜。

儘管已經有許多研究估計消費者需求，然而現有模型的估計大多不夠精確 (An et al., 2017)。(Yuan et al., 2014)提及模型選擇關鍵特徵後，使用神經網路可以使需求估計更準確。(Chen et al., 2015)使用一套整合的方法，包含資料模擬、訂價估計演算法以及長期訂價策略預測，使供應商利潤最大化。(Mostafaeipour et al., 2018)考量每個區域的收入和人口規模，使用神經網路估計航空旅行需求。(Pan et al., 2018)考量航班特徵、出發時間以及售出機票數量，透過 LSTM 模型估計航空公司消費者需求。然而這些需求估計模型大多都是用少量歷史資料做需求學習，沒有使用足夠多的歷史資料做出更準確的需求估計。

也有很多文獻考慮了歷史價格對於需求函數的影響。(Popescu et al., 2007)研究了歷史價格對於動態訂價模型的最優訂價策略的影響。(Caro & Gallien, 2007)參考歷史資料並蒐集最新資料來時刻調整未來決策。然而，在涉及具有銷售淡忘季節產品的實際環境中，通常很少或根本沒有可用的銷售歷史，加上環境不確定因素，這使得很難準確地透過需求學習來即時改變決策。

➤ 戰略性消費者行為

戰略性消費者行為(Strategic Customer Behavior)影響動態訂價則是近期研究關注主題，消費者到達時不會立即決定是否購買商品，而是等待價格是否下降，這種情況下之需求取決於過去、現在和未來的銷售價格(Su et al., 2007, Aviv et al., 2008)。其中戰略性消費者行為結合需求學習的動態訂價研究尚未成熟，且大部分的研究都沒有了解消費者支付意願之分布(Chade et al., 2002, Mersereau et al, 2012)。

➤ 競爭

也有研究將競爭對手作為動態訂價之影響，考量競爭對手之訂價動態調整自身價格(Kachani et al., 2007)，以及根據個別消費者的數據及行為模式動態調整其

價格(Priester et al., 2020)。然而此類問題無法精準估計消費者需求，造成庫存管理不易。

本研究屬於未知需求函數與有限庫存之動態訂價問題。由以上文獻回顧可知，需求學習有許多在有限時程內銷售有限庫存的問題，並且在銷售開始之前，對於市場需求狀態是未知的，且消費者支付意願分布也可能是未知的，需要透過觀察已知的消費者購買資料，精準估計消費者支付意願，並時時刻刻進行決策的優化。而現有的文獻中，歷史資料不足造成估計消費者需求不夠準確，且若是供應商希望透過需求學習即時調整最優策略，會面臨到歷史資料不足以及環境不確定因素影響，無法達到最佳訂價策略，造成供應商利潤損失。

2.2 人工智慧在動態訂價之應用

本節將介紹人工智慧在動態訂價之應用的發展背景與研究歷史，並說明現有文獻中面臨到的問題與挑戰，並回顧時間序列神經網路之研究以及相關損失函數之研究。

2.2.1 機器學習在動態訂價之應用

機器學習在近期有突破性的成長，最著名的莫過於 2016 年打敗世界棋王的 AlphaGo，其使用深度卷積神經網路以及強化學習技術，而機器學習可以對影響需求的許多現象進行建模，如競爭、未知且波動的需求或是戰略性購買行為，因此在動態訂價之應用方面有許多研究。

(Shakya et al., 2008)使用演化演算法分析動態訂價中需求不確定性的影響，然而當需求波動較低時效果較差。(Kutschinski et al., 2003)使用 Q 學習研究不同訂價策略在不同競爭水平下的市場學習行為。(Lu et al., 2018)將 Q 學習運用於智慧電網動態定價上。(Schwind et al., 2002)使用強化學習運用於資訊產品的動態定價。(Maestre et al., 2019)使用強化學習技術來解決動態定價問題。(Mullen et al., 2006)



使用粒子群最佳化(Particle Swarm Optimization)找出未知的需求參數已進行動態訂價。然而使用這些簡化之機器學習方法會導致訂價結果出現不可避免的失真。

2.2.2 神經網路在動態訂價之應用

神經網路則是機器學習其中一種強大的工具，可以協助人類做出明智的決策，因為他們可以學習和建模複雜的輸入輸出資料之間的關係，甚至求解非線性資料對他們來說也是輕而易舉。神經網路主要用於分類問題、回歸問題以及時間序列預測等問題，也有許多研究透過神經網路探討訂價問題。

(Ghose et al., 2009)透過神經網路計算產品的競爭售價，並透過模型動態調整訂價使供應商收入最大化。(Ghose et al., 2010)依據消費者購買決策制定出初始售價，並透過神經網路計算產品的競爭銷售價個。(AmalNick et al., 2017)在需求模型上創建一個神經網路，並透過進化演算法結合神經網路來優化訂價策略。(Kong et al., 2004)透過神經網路在不確定性市場上進行動態訂價策略。(Genser et al., 2022)透過多層神經網路預測訂價，達到最佳解。(Shakya et al., 2012)建立了一個基於神經網路的需求模型，並使用演化演算法來優化建立模型的策略。(Teodorović et al., 2007)結合神經網路與動態規劃來優化道路通行費。

由上述文獻可知，神經網路可以學習複雜輸入輸出資料之間的關係，因此本研究使用神經網路學習消費者需求，並結合適當動態規劃求解訂價，即可以在銷售時程開始前訂定初始訂價，並透過銷售時程的前進即時優化訂價，使供應商利潤最大化。

2.2.3 時間序列預測之神經網路

一般神經網路被標記為不適合時間資料預測，原因不僅是因為單個時間序列本身通常太短，無法使用複雜的方法進行建模，且時間序列的特徵會隨著時間的

推移而變化，即使是較長的時間序列也可能不包含足夠的相關數據來擬合複雜的模型。

遞迴神經網路(RNN, Recurrent neural networks)於 1986 年提出 (Rumelhart et al., 1986)，是序列預測中最常用的神經網路架構，在處理時間序列資料上效果顯著，其循環單元中的反饋迴路(Feedback Loops)本質上解決時間序列的依賴性問題 (Schäfer et al., 2006)。然而當時間週期過長時序列資料存在梯度消失(Gradient Vanishing)或梯度爆炸(Gradient Exploding)問題(Bengio et al., 1986 ; Pascanu et al., 2013)，這兩個棘手的梯度問題阻礙 RNN 細胞捕獲資料長期依賴性的能力。長短期記憶遞迴神經網路(Long Short-Term Memory neuron network, LSTM)於 1997 年提出(Hochreiter et al., 1997)，其神經元可以捕獲序列中的長期依賴性，同時解決 RNN 中梯度的問題(Chung et al., 2014)。GRU (Gated Recurrent Unit)則是 LSTM 的延伸，其結構更為簡單。

(Sundermeyer et al., 2015)透過 RNN 及 LSTM 估計單字序列的概率分布並進行比較，並證明 LSTM 模型表現更好。(Shewalkar et al., 2019)評估 RNN、LSTM 和 GRU 的效能，結果表示 LSTM 效力最佳，而 GRU 計算速度最快。(Cahuantzi et al., 2023)研究證明在較複雜的序列中，LSTM 表現較 GRU 出色。(Astawa et al., 2022)使用 RNN、LSTM 和 GRU 預測網站訪客數量，測試結果為使用 LSTM 預測得到之 MSE 最小。(Mahjoub et al., 2022)比較 LSTM、GRU 和 Drop-GRU 預測短期電力消耗之能力，結果顯示 LSTM 模型預測能力最好。(Abdella, 2021)指出 LSTM 非常適合用在動態訂價和需求學習之問題。

由上述文獻可以發現，LSTM 在預測時間序列資料能力突出，能解決 RNN 梯度消失或梯度爆炸問題，且當序列較為複雜時，LSTM 模型預測效果也較 GRU 更好，且 LSTM 也適合用在動態訂價和需求學習之問題。

有關長短期記憶遞迴神經網路應用於訂價方面的研究已經十分成熟。(Nelson et al., 2017)透過 LSTM 學習股票市場歷史價格及技術分析來預測股票價格未來趨



勢。LSTM 也可以估計具有時序特徵的分布資料。(Yu et al., 2019)使用 LSTM 模型來估計系統的剩餘使用壽命分布，並使用大均值差(MMD)補償損失函數以縮小分布的差異。(Sun et al., 2020)使用 LSTM 進行風電超短期概率估計，建立條件正態分布模型來描述估計誤差的不確定性。(Li et al., 2020)使用 LSTM 模型估計土壤溫度的分布，並表明其效果顯著。(Liu et al., 2018)使用 LSTM 模型估計城市地區的短期食品配送需求分佈，且需求會不斷變化造成估計需要即時更新。

根據文獻可以發現，許多研究藉由 LSTM 估計具有時序特徵的分布，因此只要給予其適當的訓練資料及適合的損失函數，就可以應用於本研究，以估計消費者支付意願分布。

2.2.4 損失函數的選擇

損失函數是用來衡量模型的估計與實際值之間的差異。選擇合適的損失函數取決於估計目標和模型的特性，使用適當的損失函數才會使估計值越來越靠近真實目標。在訂價模型中有各式各樣的損失函數選擇，(Chen, 2020)使用強化學習方法計算行動邊緣運算系統的收入最大化問題，並使用 MSE 作為損失函數。(Wang, 2021)使用最小平方法 (Least Squares)作為損失函數估計需求函數。(Keskin, 2022)使用 MAE 計算需求估計參數。然而這些研究目標僅估計單一變數，無法用於估計兩個變數甚至是分布中兩個變數的差異。

(Asghari & Cyrus, 2017)在拍賣框架下提出一個乘車共享平台之訂價模型，其中使用 Kullback–Leibler divergence(KL Divergence)作為損失函數比較真實分布與估計分布之間的差異。(Javanmard, 2020)在考慮差異化產品供應商的多產品訂價模型中，使用 KL Divergence 作為損失函數。(Li, 2023)在電動車充電站之動態訂價模型中，使用強化學習演算法得到最優訂價策略，並使用 KL Divergence 作為損失函數。(Xiao et al., 2022)透過消費者購買偏好估計消費者支付意願分布，使用 KL Divergence 作為損失函數，其中支付意願分布服從常態分布。(Asghari, 2018)

在其動態訂價模型中選擇 KL Divergence 作為損失函數，而不選擇 Jensen-Shannon divergence(JS Divergence)的原因為 KL Divergence 之不對稱性(Not Symmetric)最適合計算估計分布與真實分布差異時使用。

KL Divergence 由於其計算簡單(Kullback et al., 1951)，連續分布和離散分布均適用，非負性的特性使 KL Divergence 成為一個可以用來量化兩個分布之間差異的正數，且其非對稱性最適合用來計算估計分布與真實分布差異，被廣泛應用於各式各樣的動態訂價模型中做使用。由於本研究模型亦是估計消費者支付意願分布，因此損失函數選擇 KL Divergence 作為衡量估計分布與真實分布之間的差異，使估計能越來越準確。

2.3 小結

綜觀以上現有的文獻，本研究的貢獻如下：

- i. 由 2.1 可知，若是供應商希望透過需求學習即時調整最優策略，會面臨到歷史資料不足以及環境不確定因素影響，無法達到最佳訂價策略。因此本研究模擬大量購買資料作為訓練資料，以解決需求學習資料不足之問題。
- ii. 由 2.2.2 可知，神經網路可以學習複雜輸入輸出資料之間的關係，因此本研究使用神經網路學習消費者需求。
- iii. 由 2.2.3 可知，長短期記憶神經網路(LSTM)相較於其餘時間序列預測之神經網路具有更好的估計能力，因此本研究使用 LSTM 進行環境參數識別，學習消費者支付意願之常態分布與伽瑪分布，在需求未知的環境中能做出良好初始決策。
- iv. 由 2.2.4 可知，當神經網路估計模型需要估計分布並且與真實分布做比較時，KL Divergence 最適合作為計算兩分布之間的差異。因此本研究透過 LSTM 輸出資料為該市場分布參數值落在每一種數值的機率，並使用估計機率與實際值相乘得到期望參數值，最後透過 KL Divergence 作為損失函數計

算真實分布與估計分布之間的差異，這是目前已知中沒有文獻使用過的方法。



第三章 長短期記憶神經網路之架構



本研究將使用長短期記憶神經網路學習歷史資料，以便在真實環境應用時能藉由相似歷史資料，有效辨別消費者支付意願(WTP)分布。本章節首先介紹長短期記憶神經網路，並說明資料型態以及資料模擬方式，最後介紹本研究長短期記憶神經網路之架構。

3.1 長短期記憶神經網路之介紹

本研究使用在文獻中最為普及且常用的 Vanilla LSTM 作為學習資料的模型，是在傳統遞迴神經網路(RNN)基礎上進行改良的一種網路結構，LSTM 透過增加內部三個控制門單元:遺忘門、輸入門以及輸出門對於隱藏狀態及記憶狀態進行控制，避免 RNN 中梯度消失及梯度爆炸的問題，LSTM 內部結構圖如圖 2 所示，其中 x_t 為 input data， h_t 為 output data。

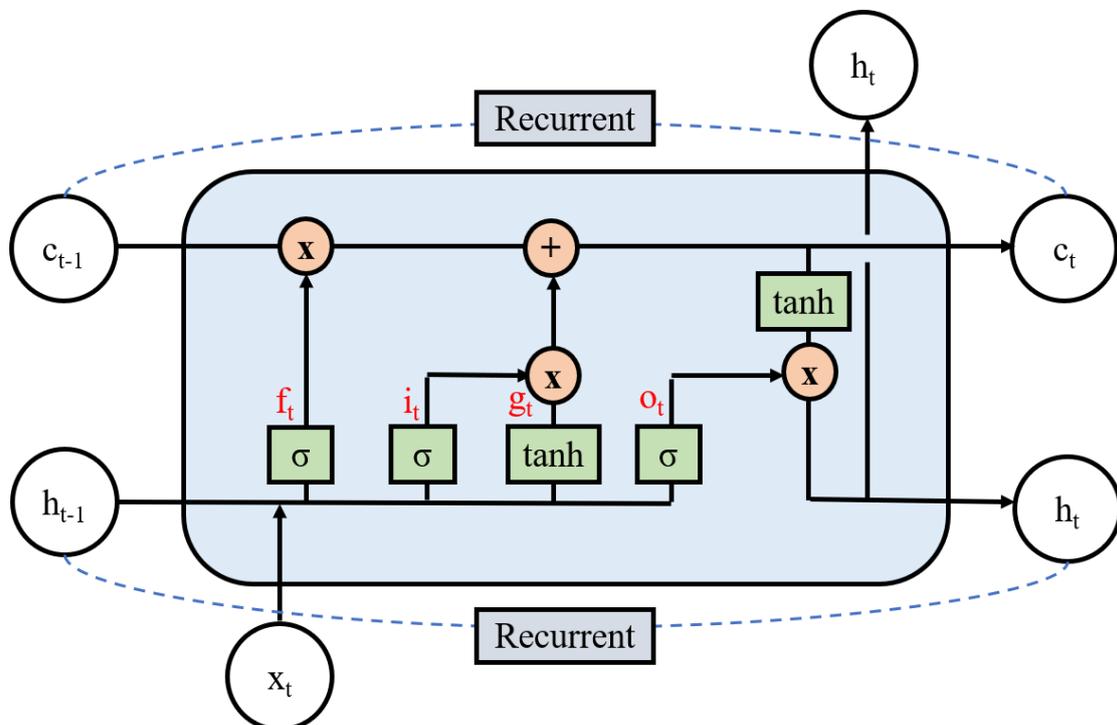


圖 2 LSTM 神經元內部結構圖



以下將對 LSTM 三個控制門單元的構造及計算進行詳細的說明:

➤ 遺忘門(Forget Gate)

決定細胞內訊息傳遞狀況，控制哪些訊息需要被丟棄，哪些需要被留存並傳遞到下一個時間，使用 sigmoid 激活函數將細胞狀態壓縮至 0 到 1 的範圍內，數值為 0 表示全部丟棄，數值為 1 則表示完全保留。

$$f_t = \sigma(W_f[h_t - 1, x_t] + b_f) \quad (1)$$

➤ 輸入門(Input Gate)

確定哪些訊息需要被丟棄，哪些可以被添加到細胞狀態內，與遺忘門一樣透過 sigmoid 激活函數將細胞狀態壓縮至 0 到 1 的範圍內，數值為 0 表示全部丟棄，數值為 1 則表示完全保留。另一方面相同的訊息透過 tanh 激活函數將細胞狀態壓縮至 -1 到 1 的範圍，相同的，數值為 1 則表示完全保留，否則全部丟棄。兩者內積運算後再加上前一期的細胞狀態即完成更新，成為這期新的細胞狀態。

$$i_t = \sigma(W_i [h_{t-1}, x_t] + b_i) \quad (2)$$

$$g_t = \tanh(W_c [h_{t-1}, x_t] + b_c) \quad (3)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot g_t \quad (4)$$

➤ 輸出門(Output Gate)

確定細胞狀態內哪些訊息需要被輸出到下一個時間，首先經過 sigmoid 激活函數來確定上一期和新輸入的訊息有多少需要被保留，結果再和透過 tanh 激活函數轉換的細胞狀態進行內積運算，這樣就確定了最終的輸出訊息。

$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \cdot \tanh(c_t) \quad (6)$$

Vanilla LSTM 透過遺忘門、輸入門及輸出門進行訊息控制，能夠決定要丟棄或保留哪些訊息，因此非常適合作為時間序列的研究工具。本研究使用上述基礎的 LSTM 架構進行適當修改，做為估計消費者支付意願分布之方法。



3.2 消費者購買資料蒐集

3.2.1 消費者購買資料模擬

因真實數據較難以取得，因此本研究將透過模擬的方式進行資料蒐集，消費者購買資料之模擬有以下幾項假設，而符號定義則參考表 1 所示：

- i. 每一期分為 M 個時間單位，每單位最多一位消費者到達，及每一時期最多 C^M 位消費者抵達。
- ii. 商品銷售時程分為 T 期，初始庫存為 N 且過程中不能補貨。
- iii. 當期消費者到達人數 C_t^R 為服從 Binomial(M, λ) 之隨機變數，其中 M 為每一期最大來客數量， λ 為單位時間消費者到達率。
- iv. 每一單位時間若有一個消費者到達，此消費者的願付價格 p_t^C 服從常態/伽瑪分布。
 - i. 每一期產生一個 $U(P^{min}, P^{max})$ 作為當期售價 p_t 。
 - ii. 售價下界 $P^{min} \leq$ 商品售價 $p_t \leq$ 售價上界 P^{max} 。
 - iii. 當消費者的願付價格 $p_t^C \geq$ 當期售價 p_t 時，此消費者購買一個商品並離開，否則此消費者會直接離開，且每位消費者最多購買一個商品。

符號	說明
T	動態規劃總期數
t	當期剩餘期數
N	初始庫存數量
s_t	t時期的狀態，代表當期剩餘庫存
A	定價範圍集合
a_t	t時期的決策，代表當期定價
p_t^C	t時期第C位顧客願付價格，從N常態/伽碼分布中隨機產生之數， $0 \leq C \leq C^M$
p^{min}	商品定價的下限
p^{max}	商品定價的上限
p_t	t時期的商品售價
λ	顧客到達率
C^M	每一期的最大來客數，也是每期單位時間數量
C_t^R	t時期的實際來客數量
n_t	t時期的實際售出數量
n_{at}	價錢為 a_t 的期望售出數量
m_t	t時期的存貨數量
l_t	t時期的顧客到達數量
$D_t(a_t)$	t時期購買商品的顧客數量
$Pr(\cdot)$	表示特定情況下的機率
$Pr(g_1 a_t)$	價錢為 a_t 時，顧客購買的期望機率
$Pr(g_0 a_t)$	價錢為 a_t 時，顧客不購買的期望機率
R_t	第t期的利潤
$V_t(S_t)$	狀態在 S_t 時，從t到T的最佳期望利潤

表 1 數學符號定義表

消費者購買資料模擬流程圖如圖 3 所示，每單位時間內沒有消費者抵達、有一個消費者抵達且購買一產品亦或是不購買直接離開，都算是完成一個時間周期 (period)，當完成資料所需模擬時間周期即進入資料特徵萃取的部分，也代表完成一次產品銷售時程 (horizon)，當完成資料所需銷售時程即結束模擬流程。

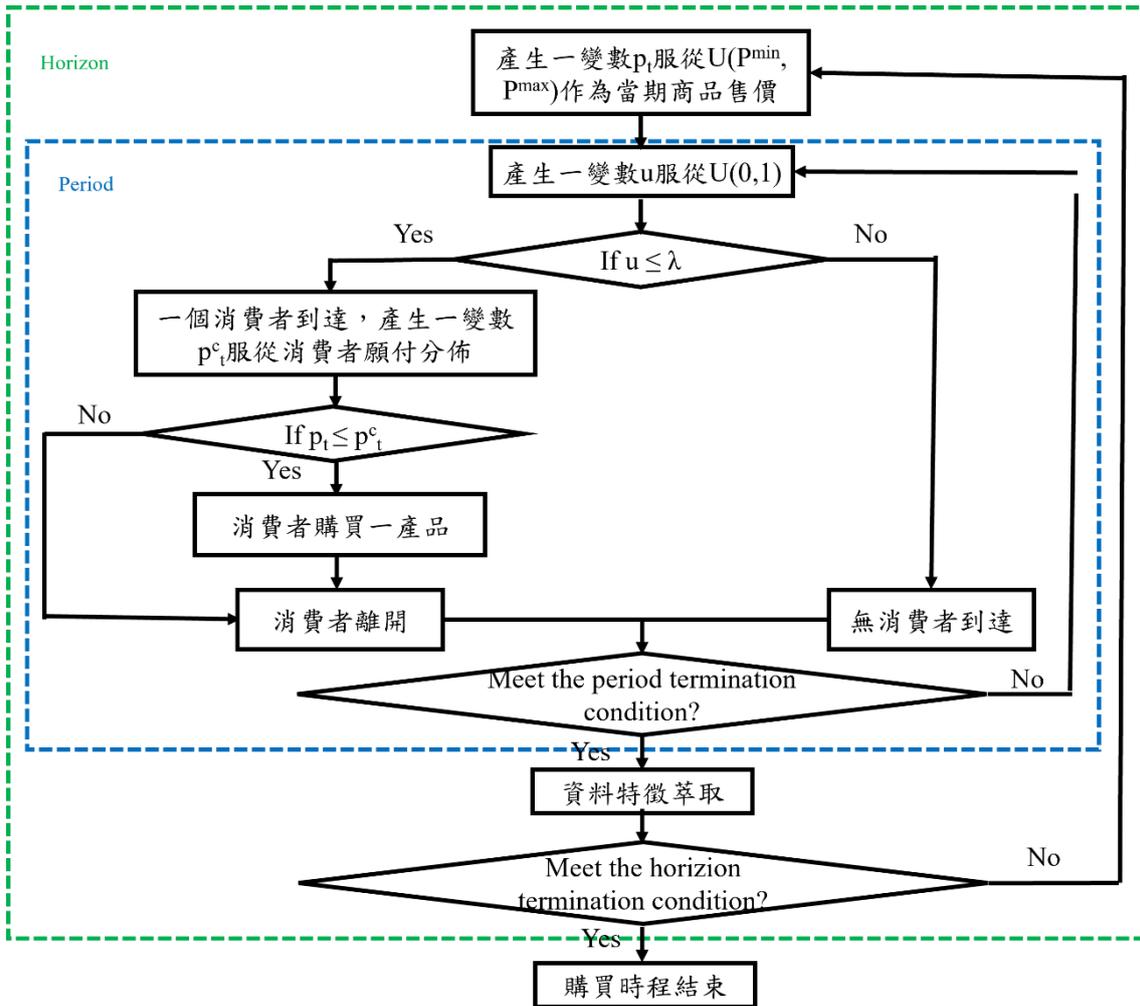


圖 3 消費者購買資料模擬流程圖

3.2.2 資料特徵萃取

➤ 輸入資料

LSTM 模型通過使用需求歷史記錄中的時間序列數據特徵來學習消費者支付意願分布。模型可以透過銷售時程不斷前進，觀察每一期的商品價格、銷售數量以及來客人數的變化，學習消費者支付意願之分布。由於 LSTM 模型需要三維度的資料:批次大小(Batch Size)、時間步長(Time Step)以及輸入特徵數(Input Size)，因此輸入資料如圖 4 所示，包含當期剩餘銷售時間 t 、前一期銷售價格 p_{t-1} 、前一期來客數 l_{t-1} 以及前一期銷售數量 n_{t-1} 四種特徵作為輸入資料，輸入資料即為圖 2 之 x_t ，每一時間步長的輸出將作為下一時間步長的輸入，使模型能捕捉時間序列資料之關聯性，使模型估計越來越準確。其中模型時間步長設置為 250，批次大小為本研究之訓練資料筆數。

本研究不對資料進行額外標準化(Standardization)或正規化(Normalization)，因為可能會損害 LSTM 估計消費者支付意願分布的能力。(DeShon, 2003)提及資料進行預處理可能會使原始數據的解釋能力降低，因此本研究使用原始資料進行訓練。

Remaining time(t)	Price(t-1)	Number of customers arrived(t-1)	Sales volume(t-1)
250	0	0	0
249	138	5	0
248	104	8	0
⋮	⋮	⋮	⋮
1	73	8	6

圖 4 單筆 LSTM 輸入資料型態

➤ 輸出資料

朱婉琳(2019)將輸出資料處理為該市場平均值落在每一種數值的機率，並使用 categorical_crossentropy 作為損失函數，然而此方法僅能估計單一變數，無法同時估計平均與標準差。因此本研究 LSTM 輸出資料在常態分布部分為該市場平

均值與標準差值落在每一種數值的機率，其中輸出資料對應為圖 2 之 h_t ，如圖 5 以 $N(125,20)$ 所示，並透過轉換將每一種數值的估計機率乘上該數值並相加總，得到期望平均值和期望標準差值，再透過 KL Divergence 作為損失函數計算真實分布與估計分布之間的差異。

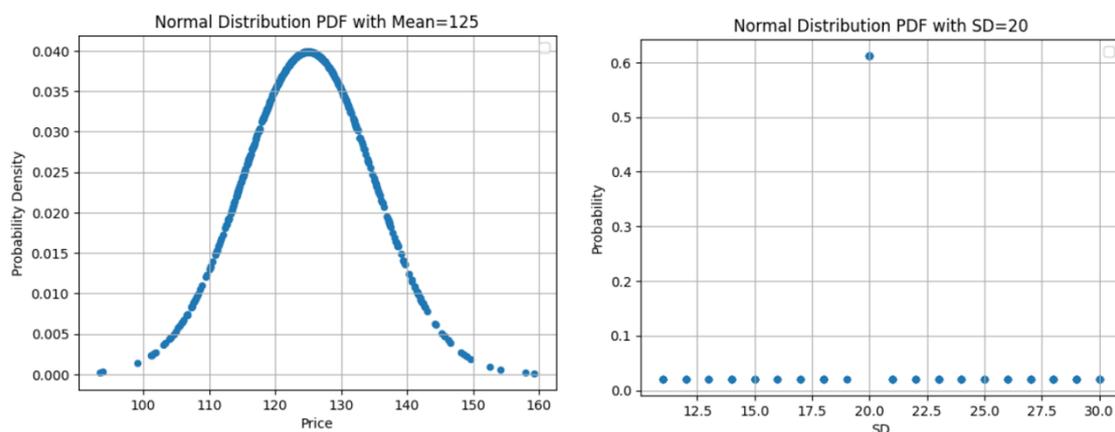


圖 5 常態市場平均值及標準差落在每一種數值的機率

本研究 LSTM 輸出資料在伽瑪分布部分為該市場形狀參數(Shape)值與尺度參數(Scale)值落在每一種數值的機率，與常態分布相同，透過轉換將每一種數值的估計機率乘上該數值並相加總，得到期望形狀參數值和期望尺度參數值，再透過 KL Divergence 作為損失函數計算真實分布與估計分布之間的差異。

3.2.3 模擬資料劃分

本研究將所有模擬資料隨機拆分，80%作為訓練資料，10%作為防止 Overfitting 的驗證資料，最後 10%作為測試資料。

3.3 神經網路架構

本研究需要每一筆資料輸入近來，就有相對應的輸出做出判斷，因此使用 (Lipton et al., 2015) 提出 Target Replication Strategy 的方法，如圖 6 所示，將模型訓練成多對多(Many-to-many)的神經網路，且每一個 LSTM 神經元皆和下一層

LSTM 神經元完全連結，即可達成本研究所需，即有一筆輸入資料就會得到一筆輸出資料，且每一個輸出都有等權重的損失計算。

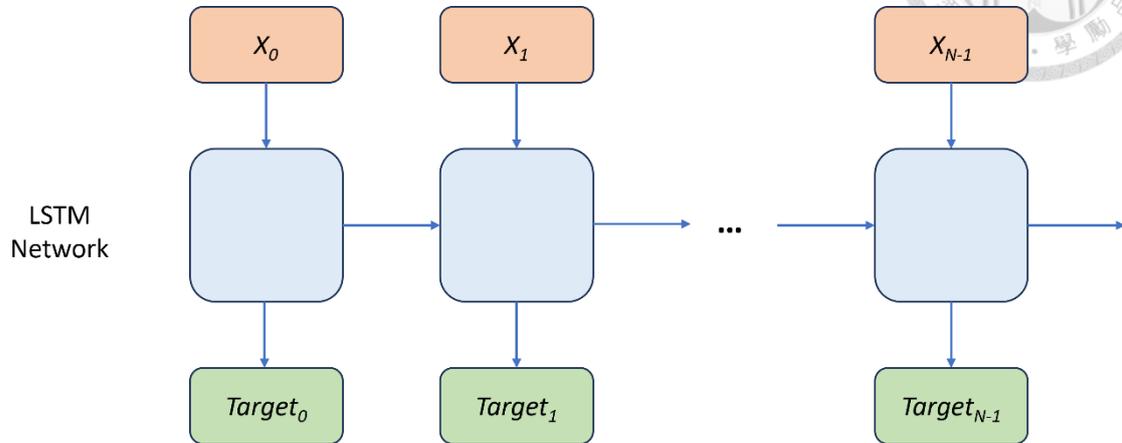


圖 6 Target Replication Strategy

LSTM 模型之超參數如果選擇不恰當，將會出現欠擬合或是過擬合的問題。因此本研究使用網格搜索(Grid Search)幫助我們找到適合的超參數。由於有太多超參數組合，當超參數太多時使用網格搜索的計算複雜度會呈現指數成長，因此本研究將由參考文獻中的訓練方式，挑選出幾種組合進行，固定其餘沒有被挑選到之超參數，最後將每一組最佳的超參數設置組合在一起，得到最適合本研究之超參數設置。

➤ Optimizer 和 Learning Rate

Optimizer 挑選三種文獻中較常出現，即 Adam、SGD 以及 RMSprop(Kumar, 2020)，而文獻中這三種的 Learning Rate 通常 SGD 選擇 0.01，Adam 及 RMSprop 通常選擇 0.001，因此我們測試 Learning Rate 0.01、0.05、0.001、0.005、0.0001 和 0.0005，因此總共有 3*6 共 18 種可能，而網格搜索結果最佳超參數為 Optimizer: Adam，Learning Rate: 0.0005。

➤ Dropout Rate 和神經元數量



Dropout Rate 挑選 0.2 到 0.8，神經元數量測試 100、200 到 600(Srivastava, 2014)，因此共有 7*6 共 42 種可能，而網格搜索結果最佳超參數為 Dropout Rate:0.2，神經元數量 500。

➤ Activation Function 和 Batch size 大小

Dense 層的 Activation Function 挑選 ReLU、LeakyReLU、ELU 和 Softplus(Nwankpa, 2018)，Batch size 挑選 64、96、128 和 192，因此共有 4*4 共 16 種可能，而網格搜索結果最佳超參數為 Activation Function: LeakyReLU，Batch size:128。

本研究之神經網路的架構圖如圖 7 所示，其中因為輸出為環境參數值落在每一種數值的機率，因此輸出層的 Activation Function 選擇 Softmax，神經網路的架構圖以估計常態分布為例：

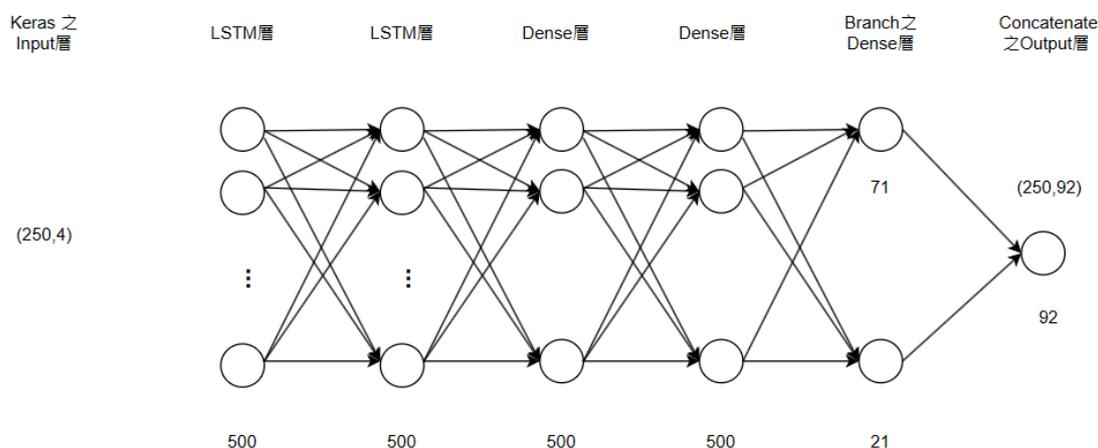


圖 7 LSTM 神經網路架構圖(常態分布)

由於本研究需同時估計分布的兩個參數，因此與先前模型不同的地方為第二層 Dense 層後面，本研究之模型將其分支成兩個 Dense 層，以分開處理兩個參數，最後再將其合併成一個 Output 層，使神經網路能同時使用兩個估計參數做損失函數之計算。



輸入之訓練資料共有 22720 筆，每筆資料有 250 時間步長，且有四個特徵。

首先透過 Kears 之 Input 層指定輸入資料的形狀，通過兩層 LSTM 層

➤ 第一層 LSTM 層

Batch Normalization 層+ Dropout 層(0.2)

➤ 第二層 LSTM 層

Batch Normalization 層+ Dropout 層(0.2)

➤ 第一層 Dense 層

TimeDistributed 層+ Batch Normalization 層+Activation Function LeakyReLU + Dropout 層(0.2)

➤ 第二層 Dense 層

TimeDistributed 層+ Batch Normalization 層+Activation Function LeakyReLU + Dropout 層(0.2)

➤ Branch 之 Dense 層

兩個參數之輸出 TimeDistributed 層+Activation Function Softmax

常態分布會 Branch 成平均以及標準差兩種 Dense 層，各有 71 種以及 21 種數值，最後將兩個輸出層 Concatenate 起來，合併成 92 種數值。伽瑪分布則是 Branch 成形狀參數與尺度參數兩種 Dense 層，各有 21 種以及 4 種數值，最後 Concatenate 起來，合併成 25 種數值。其中 Dense 層應用於 TimeDistributed 架構下，使每個時間步長的輸入資料能獨立地被處理，並在整個時間序列中學習。

至於 Loss 使用 KL Divergence 作為神經網路的損失函數，計算估計分布與真實分布之間的差異，常態分布 KL Divergence 計算方式如下(Ji et al., 2018)，其中真實分布 $N_1(\mu_1, \sigma_1)$ ，估計分布 $N_2(\mu_2, \sigma_2)$:

$$KL Divergence = \frac{1}{2} \left(\frac{\sigma_1^2 + (\mu_1 - \mu_2)^2}{\sigma_2^2} - \ln \frac{\sigma_1}{\sigma_2} - 1 \right) \quad (7)$$

伽瑪分布計算方式如下(Delpha et al., 2017)，其中真實分布 $\Gamma_1(\alpha_1, \theta_1)$ ，估計分布 $\Gamma_2(\alpha_2, \theta_2)$ ， ψ 為 digamma function:

$$KL Divergence = (\alpha_1 - \alpha_2)\psi(\alpha_1) - \log\Gamma(\alpha_1) + \log\Gamma(\alpha_2) + \alpha_2(\log\theta_2 - \log\theta_1) + \alpha_1 \frac{(\theta_1 - \theta_2)}{\theta_2} \quad (8)$$

由於 KL Divergence 計算需要使用實際平均值以及標準差值，或是實際形狀參數與尺度參數做計算，因此模型輸出需要透過轉換才能使用 KL Divergence 計算分布差異。常態分布需將估計平均值範圍 70-140，共 71 種數值的機率乘上該數值並相加得到期望平均值，將估計標準差值範圍 10-30，共 21 種數值的機率乘上該數值並相加得到期望標準差值，即可將期望平均值及標準差值作為估計分布的參數與真實分布做 KL Divergence 計算，計算每一步的 Loss 值並讓模型估計能越來越接近實際值。至於伽瑪分布需將估計形狀參數值範圍 10-30，共 21 種數值的機率乘上該數值並相加得到期望形狀參數值，將估計尺度參數值範圍 2-5，共 4 種數值的機率乘上該數值並相加得到期望尺度參數值，即可做 KL Divergence 計算，計算每一步的 Loss 值並讓模型估計能越來越接近實際值。

第四章 動態訂價模型



上一章節長短期記憶神經網路所識別的市場環境，將作為本章節動態訂價模型的輸入並產生最優訂價策略。本章節首先說明模型基本架構與相關假設，再來介紹動態訂價模型，並介紹動態規劃程式執行方式。

4.1 模型基本架構與假設

本研究透過 LSTM 學習消費者偏好後，使用 MDP(Markov Decision Processes) 為基礎建立訂價模型，生成最優訂價策略以達到利潤最大化，為了簡化複雜度，將連續時間切分成更小的離散時間單位，再透過動態規劃進行訂價模型建構及求解，以制定出最優訂價。以下將介紹訂價模型中相關基本假設及符號定義：

➤ 基本假設

- i. 每一期分為 M 個時間單位，每單位最多一位消費者到達，及每一時期最多 C^M 位消費者抵達。
- ii. 商品銷售時程分為 T 期，初始庫存為 N 且過程中不能補貨。
- iii. 當期消費者到達人數 C_t^R 為服從 Binomial(M, λ) 之隨機變數，其中 M 為每一期最大來客數量， λ 為單位時間消費者到達率。
- iv. 售價下界 $P^{min} \leq$ 商品售價 $p_t \leq$ 售價上界 P^{max} 。

➤ 數學符號定義

基於上述模型基本假設，透過動態規劃建立訂價模型求解出最優訂價，模型中使用之數學符號定義 3.2.1 表 1 所示。



4.2 動態訂價模型

本研究探討航空公司之機票訂價問題，主要目的為最大化航空公司的利潤，藉由前一章節所輸出的市場環境參數，結合當期所剩下的剩餘期數以及剩餘庫存，求解出當期最佳訂價。而當期的最佳訂價策略僅會影響到下一期的剩餘庫存，因此每一期的狀態僅會與上一期的決策有關，剩餘庫存的變化符合馬可夫過程中的無記憶性質(Memoryless)。

在計算動態規劃中狀態轉移機率前，要先固定市場環境的參數，計算在每一個價錢下的購買機率與不購買機率：

$$Pr(g_1|a_t) = 1 - Pr(X \leq a_t) \quad (9)$$

$$Pr(g_0|a_t) = 1 - Pr(g_1|a_t) \quad (10)$$

以上為通式計算方式，而本研究為常態分布以及伽瑪分布，因此將購買機率 $Pr(g_1|a_t)$ 改寫為常態分布以及伽瑪分布的累積機率函數公式，而常態分布原本公式積分範圍為負無限大至無限大，調整為從 0 開始積分至無限大，用意為價格不得為負數，因此從 0 開始積分，不購買機率 $Pr(g_0|a_t)$ 算式皆與 Eq.(10) 相同。

常態分布：

$$Pr(g_1|a_t) = 1 - \left(\frac{1}{\sigma\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{(a_t-\mu)^2}{2\sigma^2}} \right) \quad (11)$$

伽瑪分布：

$$Pr(g_1|a_t) = 1 - \left(\int_0^{\infty} \frac{a_t^{(\alpha-1)} e^{-\frac{a_t}{\theta}}}{\theta^\alpha \Gamma(\alpha)} \right) \quad (12)$$

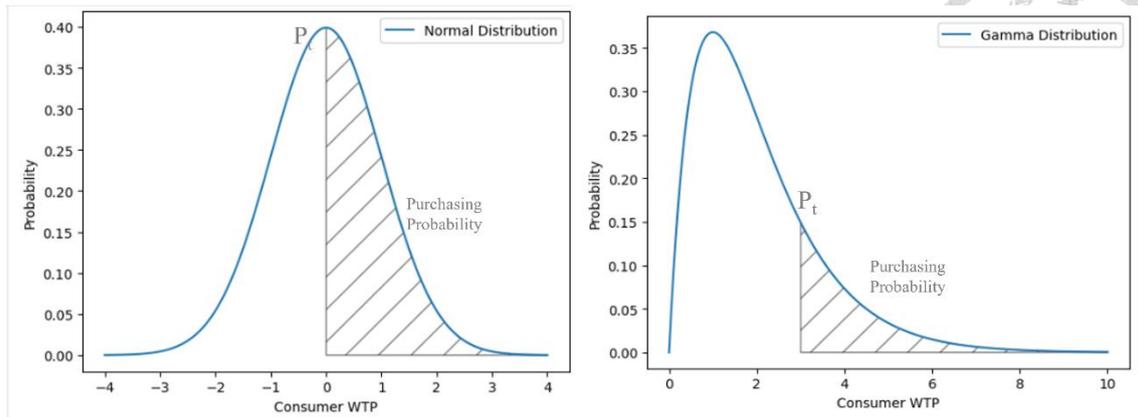


圖 8 購買機率計算示意圖

動態訂價模型將由上一小節定義的數學符號建構，以下將介紹隨機動態規劃的五大元素：決策時程(Planning Horizon)、狀態(State)、決策(Action)、狀態轉移機率(Transition Probability)及回饋函數(Reward Function)。

➤ 決策時程(Planning Horizon)

動態訂價中決策的總期數，在本研究中代表機票販售總期數，且時程會在飛機起飛之時結束，剩餘機票庫存也無價值。

$$t \in \{1, 2, \dots, T\}$$

本研究中每個時期切分成 M 個時間單位，如此連續時間將分割成離散之時間，且每個時間單位內有無消費者到達皆視為一個事件的發生。

➤ 狀態(State)

每一時期航空公司尚未賣出的機票數量，庫存為飛機座位因此不可補貨。

$$s_t \in \{1, 2, \dots, N\} \text{ and } s_T = N$$

➤ 決策(Action)

在第 t 期所做出的訂價決策，並且受到訂價範圍限制。

$$a_t \in A = \{P^{min}, P^{min} + 1, \dots, P^{max}\}$$

➤ 狀態轉移機率(Transition Probability)

每一期之間狀態轉移的機率函數，前一期的狀態與決策會影響到下一期的初始狀態。而前面有提及每一時期又分為 C^M 個時間單位，意即每一期最大來客數為 C^M 位消費者，因此在 $s_t \geq C^M$ 情況下， s_{t-1} 的可能性有 C^M 種，等同於可能賣掉 C^M 個產品，否則可能性為 s_t 種，因為僅剩下 s_t 個產品，最多只能銷售 s_t 個。

每一期最大來客數量 C^M 個，實際來的消費者為 C_t^R 個之機率：

$$Pr(C_t^R | C^M) = C_{C_t^R}^{C^M} \lambda^{C_t^R} (1 - \lambda)^{C^M - C_t^R}, 0 \leq C_t^R \leq C^M \quad (13)$$

實際來的消費者為 C_t^R 個的情況下，在定價為 a_t 時，售出 n 個產品的機率：

$$Pr(D_t(a_t) = n | C_t^R) = C_n^{C_t^R} Pr(g_1 | a_t)^n Pr(g_0 | a_t)^{C_t^R - n}, n \in [0, \min\{s_t, C_t^R\}] \quad (14)$$

得到在 a_t 訂價下售出 n 個產品的機率：

$$Pr(D_t(a_t) = n) = \sum_{C_t^R = n}^{C^M} Pr(C_t^R | C^M) Pr(D_t(a_t) = n | C_t^R) \quad (15)$$

$D_t(a_t)$ 之期望值即為：

$$n_{a_t} = E[\min\{D_t(a_t), s_t\}] \quad (16)$$

而在 a_t 訂價下售出 n_{a_t} 個產品的轉移機率：

$$s_{t-1} = \text{Max}\{0, s_t - D_t(a_t)\} \quad (17)$$

➤ 回饋函數(Reward Function)

計算航空公司在整體決策時程可以獲得的回饋，包含當期立即回饋以及未來預期回饋。

當期立即回饋為 s_t 轉變為 s_{t-1} 期時航空公司可以立即得到的回饋，在本研究中即為當期訂價乘上當期賣出的機票數量：

$$R_t(s_t, a_t) = a_t E[\min\{D_t(a_t), s_t\}] \quad (18)$$

未來預期回饋為第 t 期時，狀態 s_t 時做出 a_t 決策，預期未來可以獲得的回饋：

$$E[V_{t-1}(s_{t-1} | s_t, a_t)] \quad (19)$$



將 t 時期的立即回饋與未來預期回饋相加總，即為貝爾曼方程式，亦稱為 Optimality Equation，表示為：

$$V_t(s_t) = \max_{a_t \in A} \{R_t(s_t, a_t) + E[V_{t-1}(s_{t-1} | s_t, a_t)]\} \quad (20)$$

藉由回饋函數計算出每一種決策 a_t 下的最佳期望利潤，最後選擇出當期最佳決策，即當期最優訂價 a_t 。

最後說明動態規劃之執行流程。本研究透過 C# 撰寫，並將相關參數：當期環境參數、剩餘銷售時間、當期剩餘庫存、當期最佳訂價以及當期最佳利潤以 CSV 檔案作為輸出。

計算開始時會根據市場環境之參數，使用 Eq.(10)、Eq. (11) 及 Eq. (12) 計算每個價錢下的購買機率 $Pr(g_1 | a_t)$ 與不購買機率 $Pr(g_0 | a_t)$ ，就可以得到每個訂價策略下售出 n 個產品的機率。從剩餘時間 T 期開始販售 N 個商品，計算出售價下界 P^{min} 至售價上界 P^{max} 每個價錢下的期望機率、當期立即回饋與下一期會轉移到的狀態加上該狀態的最佳期望回饋，直到時程結束，透過遞迴的方式計算在每一個狀態下的最佳決策。

第五章 模型測試與模擬驗證



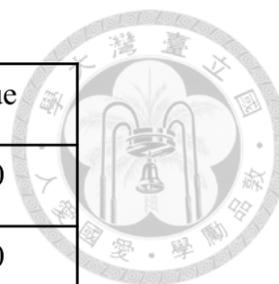
本章節會對長短期記憶神經網路進行模型測試，展示訓練成果，並比較模擬驗證得到之模擬利潤與動態規劃求解的真實利潤的差異，並計算出偏好學習成本 Gap。第一小節說明消費者支付意願服從常態分布之參數設置，長短期記憶神經網路模型測試，模擬驗證流程與模擬驗證之結果，第二小節說明消費者支付意願服從伽瑪分布之參數設置，長短期記憶神經網路模型測試，模擬驗證流程與模擬驗證之結果。

5.1 WTP 服從常態分布

5.1.1 參數設置

本小節將會對於 3.2.1 中 LSTM 消費者購買資料模擬的參數設置進行介紹，其中消費者支付意願分布服從常態分布，參數之設置如表 2 所示。

本研究中的消費者購買資料模擬中消費者支付意願服從常態分布，其中商品售價是依照常態分布之參數範圍設置，太高之價格無人購買，太低的價格皆會購買，導致成為無效資料，而太多無效資料會造成 LSTM 無法充分學習與訓練，使估計不夠準確。



Parameters	Value
Planning Horizon	250
Number of Products	300
Arrival Probability	0.7
Mean of WTP	70~140
Standard Deviation of WTP	10-30
Range of Price	50~200
Price Difference	1
Maximum Number of Customers	20

表 2 WTP 為常態分布之參數設置

5.1.2 長短期記憶神經網路模型測試

本小節將使用 3.2.3 中提及，將模擬資料中的 10%作為測試資料對 LSTM 模型進行測試，驗證模型估計準確性。

本研究中決策時程共有 250 期，測試資料共有 2840 筆，因此 LSTM 模型會對這 2840 筆資料分別進行 250 次的環境估計。圖 9 及圖 10 為當期銷售期數為 1、3、5、10、25、50、100、150 與 250 期時 LSTM 模型估計的結果，其中橫軸為估計環境平均值和標準差值，縱軸為真實環境平均值和標準差值，紅色斜虛線為標準答案，每一個藍色點代表其中一筆測試資料在當期時 LSTM 模型估計之值，越靠近紅色斜虛線代表估計越準確。由圖 9 可以觀察到在完全沒有參考數據的第 1 期估計會落在平均值範圍 70-140 的中心位置附近，並且在 25 期時估計就已經非常接近真實值，而第 250 期可以看到估計非常準確。由圖 10 可以觀察到在完全沒有參考數據的第 1 期估計沒有落在標準差值範圍 10-30 的中心位置附



近，而是落在 24 左右，並且在 25 期時估計就已經非常接近真實值，而第 250 期可以看到估計十分準確。

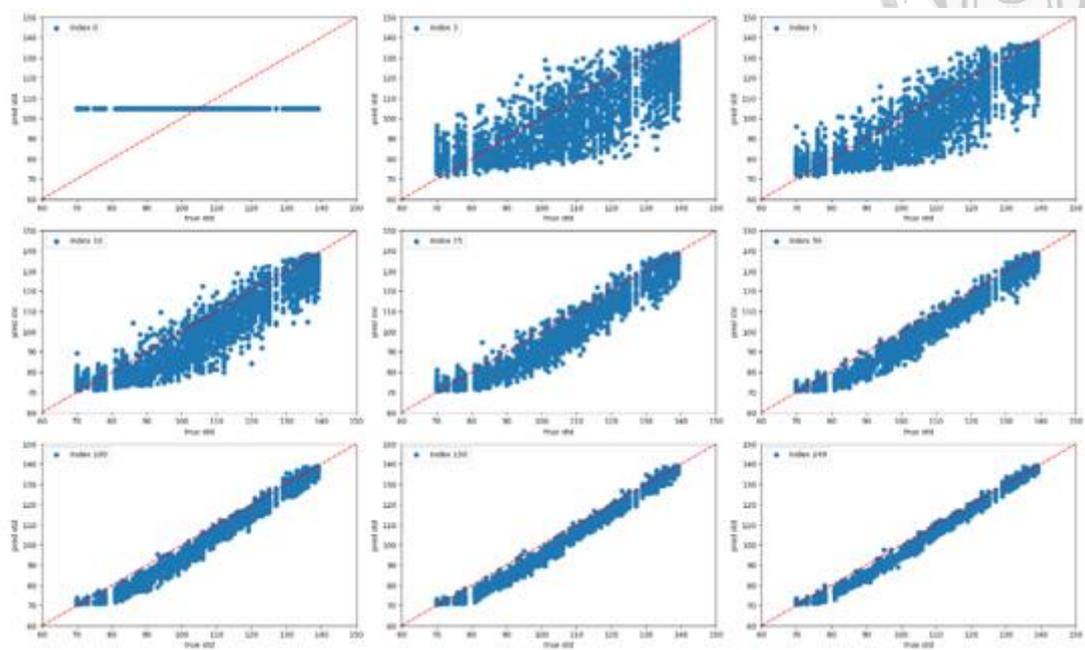


圖 9 估計 250 期銷售時程之常態分布平均值

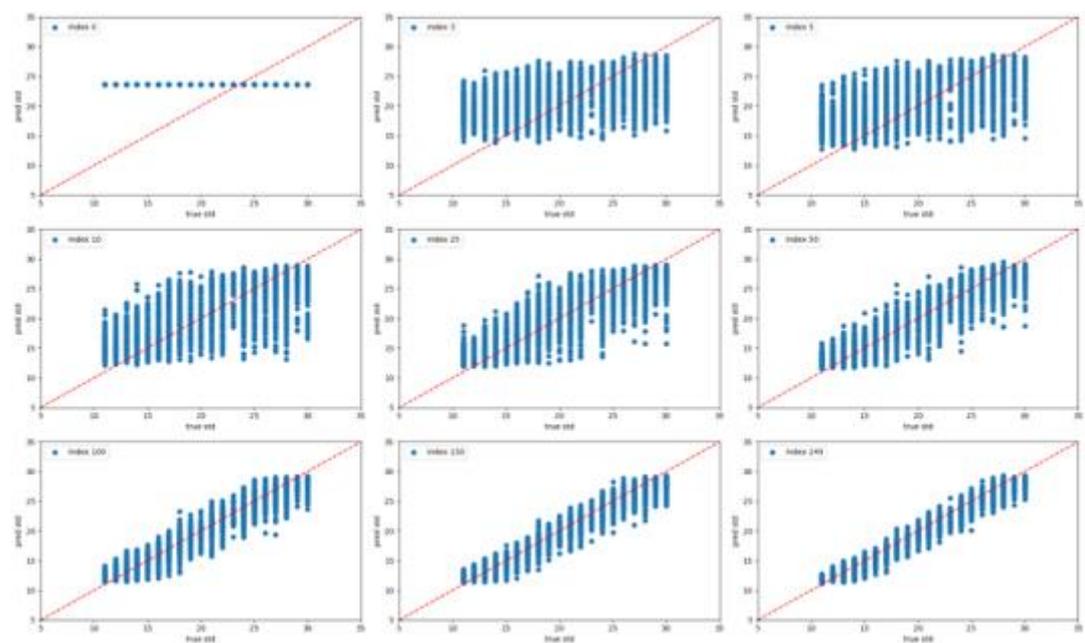


圖 10 估計 250 期銷售時程之常態分布標準差值

接下來探討為何估計之標準差值在第一期時，不是落在標準差值範圍 10-30 的中心位置 20 附近而是落在 24 左右。如表 3 所示，真實平均為 70，標準差為



20，當估計平均也為 70 而標準差為 15 和 25 時，KL Divergence 值為 0.101 及 0.043，可以發現標準差較大時 KL Divergence 值較小。另外當估計平均皆為 140，而估計標準差為 11 和 29 時，KL Divergence 值為 0.555 及 0.109，可以發現標準差較大時 KL Divergence 值也較小，因此推測模型估計第一期之標準差值會較大，是因為模型想要有較低的 Loss 值，當標準差較大時 KL Divergence 值會較小，因此第一期估計會較中心大一些。

KL Divergence	mean_true=70 std_true=20
mean_pred=140 std_pred=11	0.555
mean_pred=70 std_pred=15	0.101
mean_pred=70 std_pred=20	0
mean_pred=70 std_pred=25	0.043
mean_pred=140 std_pred=29	0.109

表 3 KL Divergence 值比較表

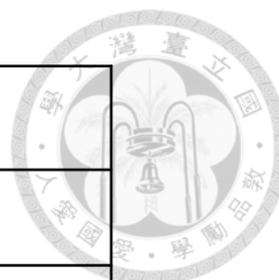
5.1.3 模擬驗證流程

為了驗證模型的效力與通用性，我們使用偏好學習成本 Gap 作為衡量標準，計算本研究模型學習真實利潤所需要的成本， \hat{R} 為在環境參數未知情況下，使用 LSTM 識別市場環境參數並使用估計參數進行動態規劃求解得到之模擬平均利

潤， R 為在已知環境參數下使用動態規劃求解得到之最佳利潤， Gap 的計算方式如下：

$$Gap = \frac{(\hat{R} - R)}{R} * 100\% \quad (21)$$

本研究將使用離散事件模擬，其中常態分佈模擬平均範圍 70-140，模擬標準差選擇範圍下界、範圍中心以及範圍上界 10、20、30 作為模擬標準差，總共 213 種環境。伽瑪分布則是模擬形狀參數範圍 10-30，尺度參數 2-5 總共 84 種環境。每一種環境皆模擬 30 次，以下將進行模擬流程之說明，其中模擬過程中會用到之參數如表 4 所示。



符號	說明
p_t	t時期的商品定價
p_t^C	t時期第C位顧客願付價格，從 $N(\mu, \sigma)$ 中隨機產生之數， $0 \leq C \leq C^M$
r_t	t時期的模擬利潤， $r_t = Q_t * n_t$
n_t	t時期的實際售出數量
m_t	t時期的存貨數量
l_t	t時期的顧客到達數量
B	用已知平均 μ 及標準差 σ 做動態定價決策得到之最佳利潤
\hat{B}	模擬總利潤， 為250期 r_t 相加總
$N(\mu, \sigma)$	顧客願付價格分布，服從平均及標準差皆為已知之常態分佈
$\hat{N}(\hat{\mu}, \hat{\sigma})$	顧客願付價格分布，服從平均及標準差皆為未知之常態分佈
$\Gamma(\alpha, \theta)$	顧客願付價格分布，服從形狀參數及尺度參數皆為已知之伽瑪分佈
$\hat{\Gamma}(\hat{\alpha}, \hat{\theta})$	顧客願付價格分布，服從形狀參數及尺度參數皆為未知之伽瑪分佈
λ	顧客到達率
C^M	每一期的最大來客數，也是每期單位時間數量
\hat{R}	模擬平均利潤
R	真實平均利潤

表 4 模擬驗證參數表

- 
- i. 在未知的環境市場中使用 LSTM 模型，進行環境參數識別，得到估計之常態分佈參數 $\hat{\mu}$ 和 $\hat{\sigma}$ ，並使用估計之環境參數得到當期售價 p_t 。
 - ii. 使用已知環境參數 μ 和 σ 、當期售價 p_t 、消費者到達率 λ 、最大來客數 C^M 以及目前剩餘存貨 m_t 進行商品銷售模擬。
 - iii. 每一單位時間若有消費者到達，此消費者的消費者願付價格 p_t^C 服從常態分布。
 - iv. 當消費者的願付價格 $p_t^C \geq$ 當期售價 p_t 時，此消費者購買一個商品並離開，否則此消費者會直接離開，且每位消費者最多購買一個商品。
 - v. 當商品銷售完或是售出數量 $n_t <$ 存貨數量 m_t 時，則視為購買結束，此時當期模擬利潤 $r_t =$ 商品售出數量 $n_t *$ 當期售價 p_t 。
 - vi. 紀錄當期售價 p_t 、消費者到達數量 l_t 、商品售出數量 n_t 作為下一期之 LSTM Input。
 - vii. $\hat{B} = \sum r_t$ 。
 - viii. B 為每次模擬時，用已知環境參數 μ 和 σ 做動態規劃訂價求解得到之最佳真實利潤。
 - ix. $R = \frac{\sum \hat{B}}{\text{模擬次數 } 30}$ 。
 - x. $\hat{R} = \frac{\sum B}{\text{模擬次數 } 30}$ 。
 - xi. 計算 Gap，得到學習真實利潤需要的成本。

5.1.4 模擬驗證結果

本研究常態分布模擬平均值範圍為 70-140，模擬標準差值為 10、20、30。本小節將詳細介紹模擬驗證結果，並使用模擬過程中的各種輔助圖，證明偏好學習成本 Gap 結果無誤。

➤ 標準差為 10 之模擬驗證結果

如 5.1.3 提及之流程，首先會使用 LSTM 模型估計環境參數，得到估計之 $\hat{\mu}$ 和 $\hat{\sigma}$ ，因此將估計之 $\hat{\mu}$ 和 $\hat{\sigma}$ 計算常態分布 KL Divergence，得到結果為 0.1027，符合訓練 LSTM 模型時模型收斂之對應 Loss 值，並在當期銷售期數為 1、3、5、10、25、50、100、150 與 250 期時 LSTM 模型估計 $\hat{\mu}$ 和 $\hat{\sigma}$ 的結果做成圖 12，證明 LSTM 模型之估計效力良好。

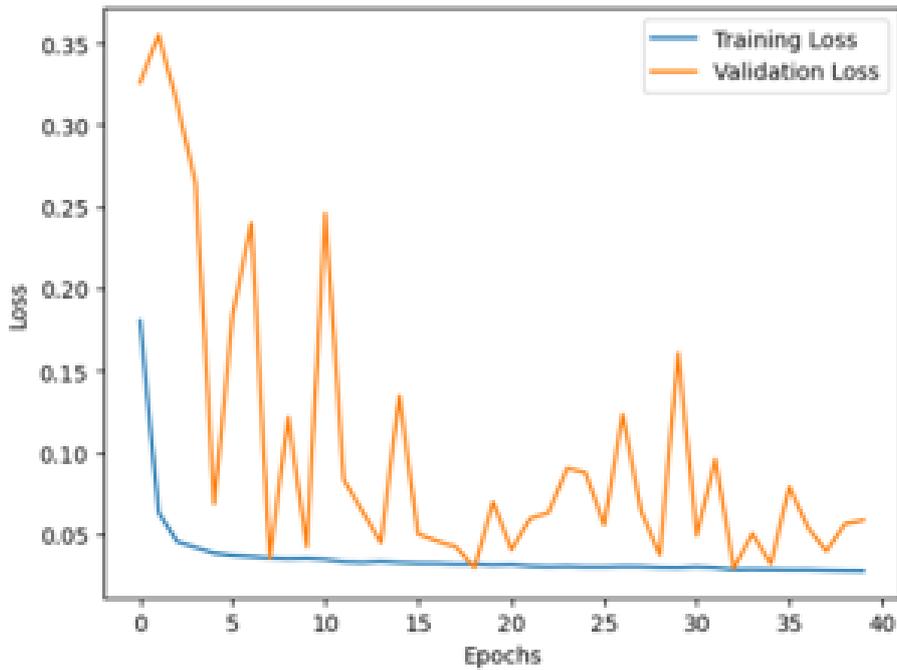


圖 11 LSTM 模型之訓練常態分布 Loss 圖

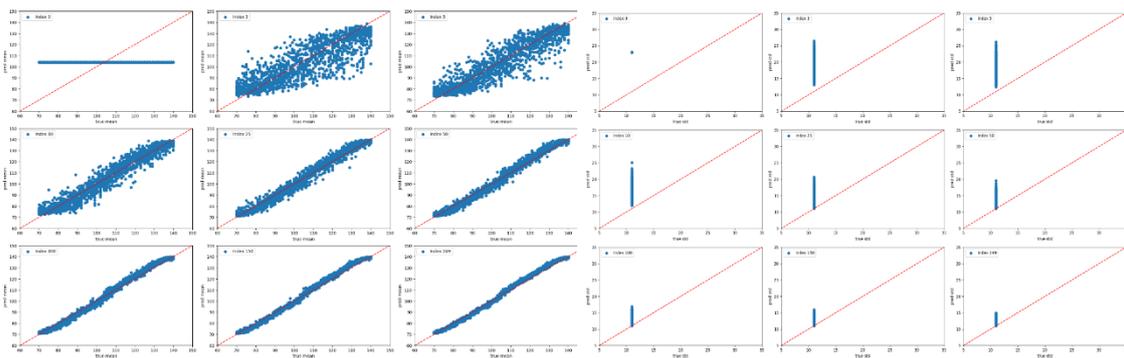


圖 12 模擬標準差為 10 之估計 250 期銷售時程的 $\hat{\mu}$ 和 $\hat{\sigma}$



接者在商品銷售模擬中，消費者願付價格 p_t^C 會服從 $N(70,10)$ ，因此我們將 30 次模擬中的第一次模擬，過程中所有產生的消費者願付價格 p_t^C 轉換成機率密度函數柱狀圖，證明 p_t^C 之正確性與真實分布相符。



圖 13 模擬標準差為 10 之第一次模擬之消費者願付價格機率密度函數柱狀圖

再者我們將全部 71 種平均值模擬 30 次之利潤做分析。如圖 14 所示，藍點為第一次模擬，平均 70-140 總共 71 種模擬利潤，紅色虛線為動態規劃求解之真實利潤，可以得知隨著平均值增加，模擬利潤以及真實利潤皆穩定上升。圖 15 藍點為平均為 70 之 30 次模擬利潤，紅色虛線為動態規劃求解之真實利潤，可以得知相同平均值情況下，30 次模擬之利潤差異不大，證明本研究模擬之可靠性及正確性。

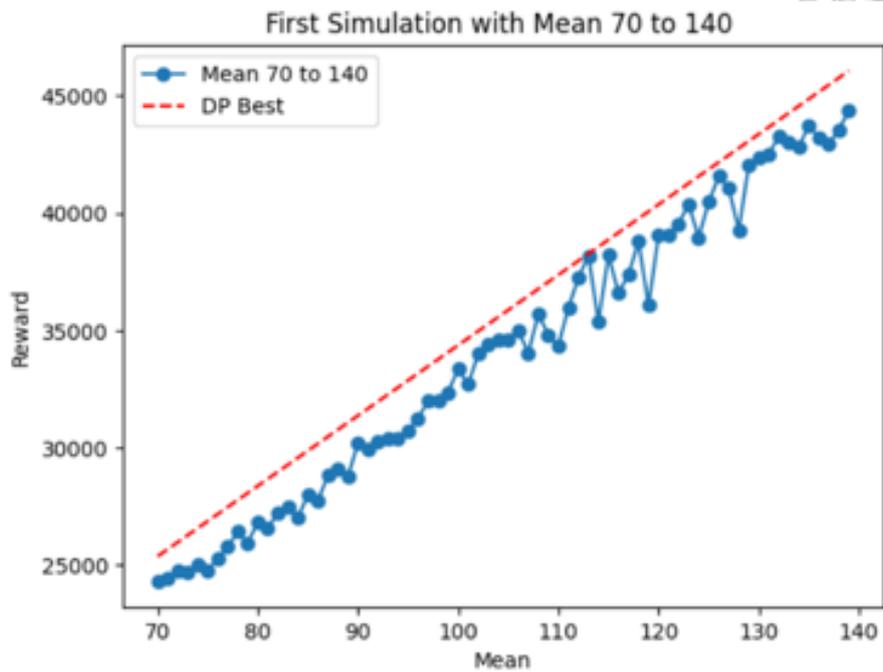
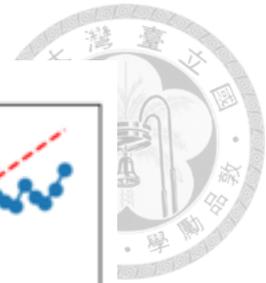


圖 14 模擬標準差為 10 之第一次模擬 71 種平均之模擬利潤

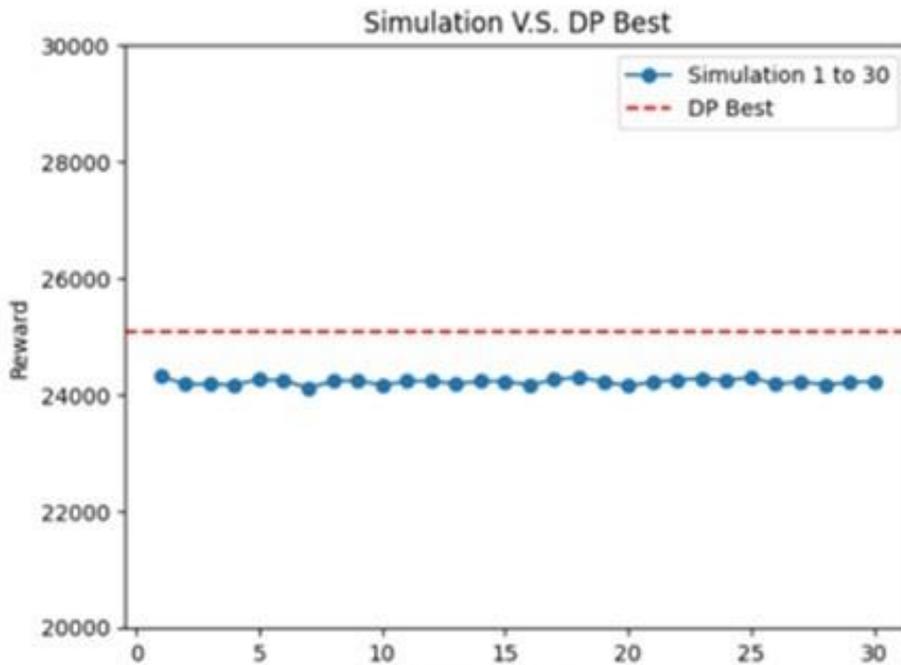


圖 15 模擬標準差為 10 之平均為 70 的 30 次模擬利潤

最後我們在已知環境參數的情況下做模擬訂價，並與動態規劃求解答案計算偏好學習成本 Gap，即為圖 16 之灰點，得到 Gap 為-0.17%，說明用已知的 μ 和 σ 做模擬訂價得到的利潤，與真實利潤差異甚小，證明整體模擬驗證流程及方法之

正確性。圖 16 之藍點則為標準差為 10 之最終模擬驗證結果，Average Gap 為-3.96%，說明本研究之方法僅需要少許的成本，即可學習真實的利潤。另外黑色虛線是動態規劃求解之真實利潤，紅色實線是模擬 30 次之平均利潤。

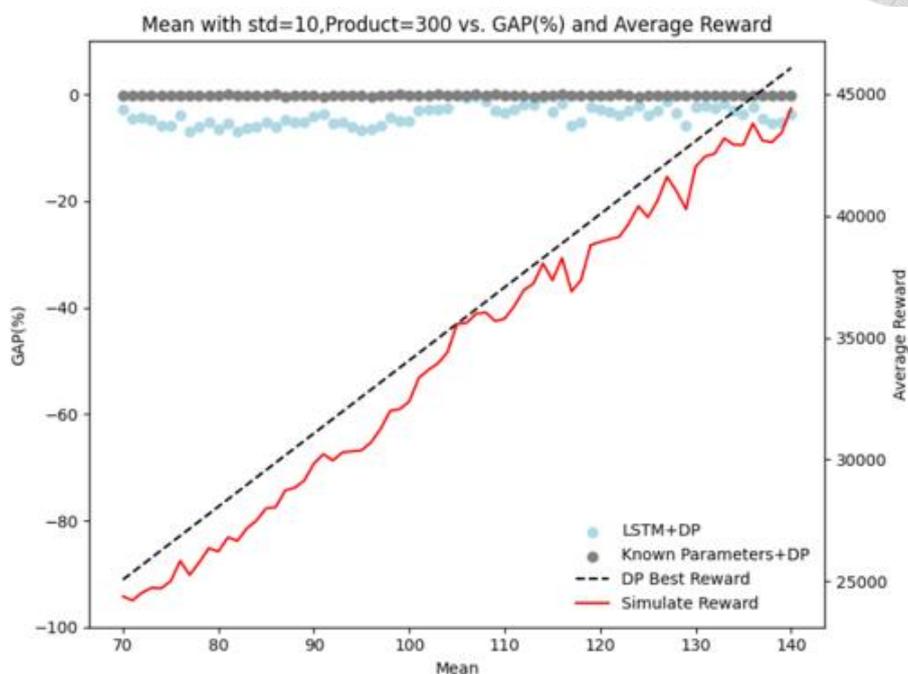


圖 16 模擬標準差為 10 之偏好學習成本圖

➤ 標準差為 20 之模擬驗證結果

將估計之 $\hat{\mu}$ 和 $\hat{\sigma}$ 計算常態分布 KL Divergence，得到結果為 0.0732，符合訓練 LSTM 模型時模型收斂之對應 Loss 值，並在當期銷售期數為 1、3、5、10、25、50、100、150 與 250 期時 LSTM 模型估計 $\hat{\mu}$ 和 $\hat{\sigma}$ 的結果做成圖 17，證明 LSTM 模型之估計效力良好。

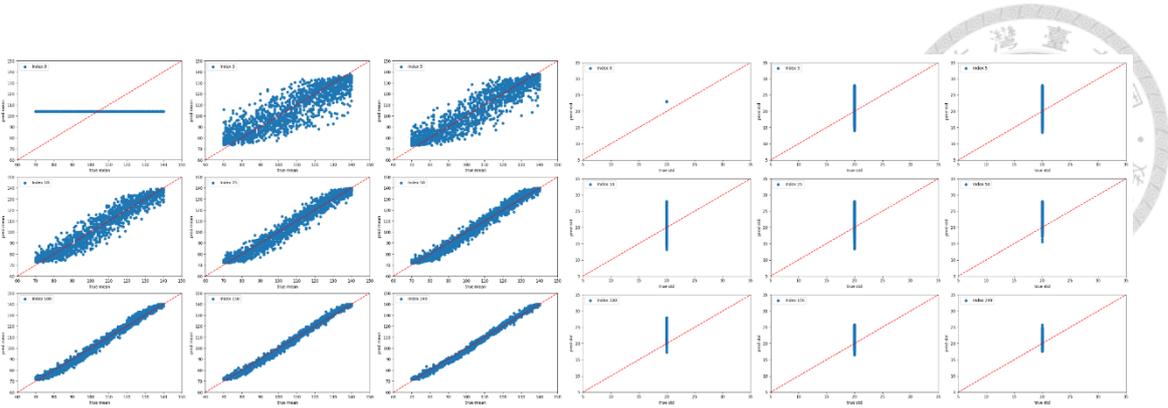


圖 17 模擬標準差為 20 之估計 250 期銷售時程的 $\hat{\mu}$ 和 $\hat{\sigma}$

接者在商品銷售模擬中，消費者願付價格 p_t^c 會服從 $N(70,20)$ ，因此我們將 30 次模擬中的第一次模擬，過程中所有產生的消費者願付價格 p_t^c 轉換成機率密度函數柱狀圖，證明 p_t^c 之正確性與真實分布相符。



圖 18 模擬標準差為 20 之第一次模擬之消費者願付價格機率密度函數柱狀圖

再者我們將全部 71 種平均值模擬 30 次之利潤做分析。如圖 19 所示，可以得知隨著平均值增加，模擬利潤以及真實利潤皆穩定上升。由圖 20 可以得知相同平均值情況下，30 次模擬之利潤差異不大，證明本研究模擬之可靠性及正確性。

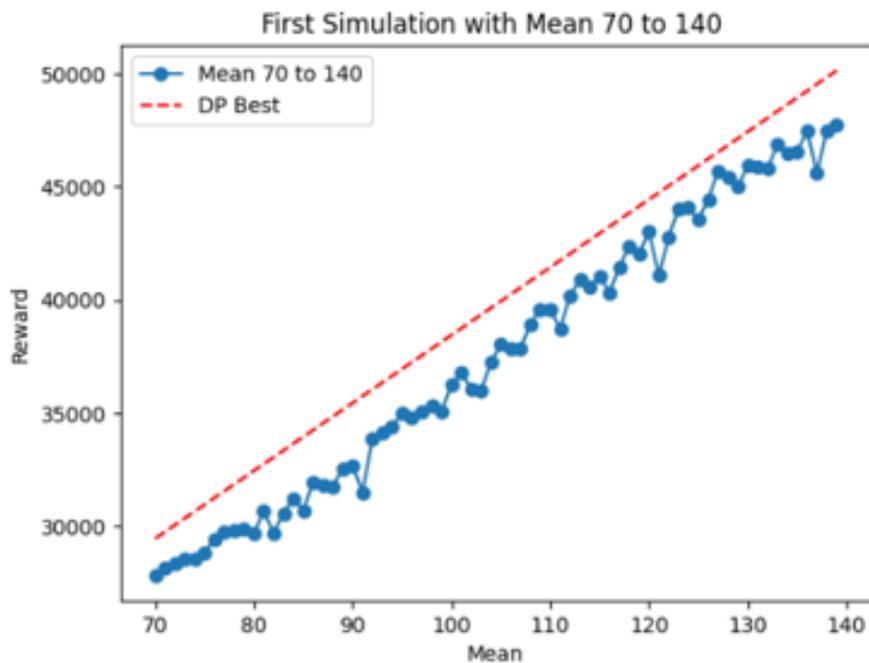
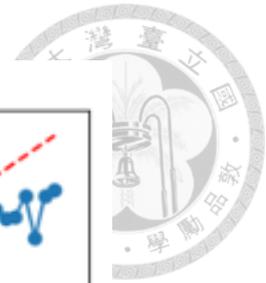


圖 19 模擬標準差為 20 之第一次模擬 71 種平均之模擬利潤

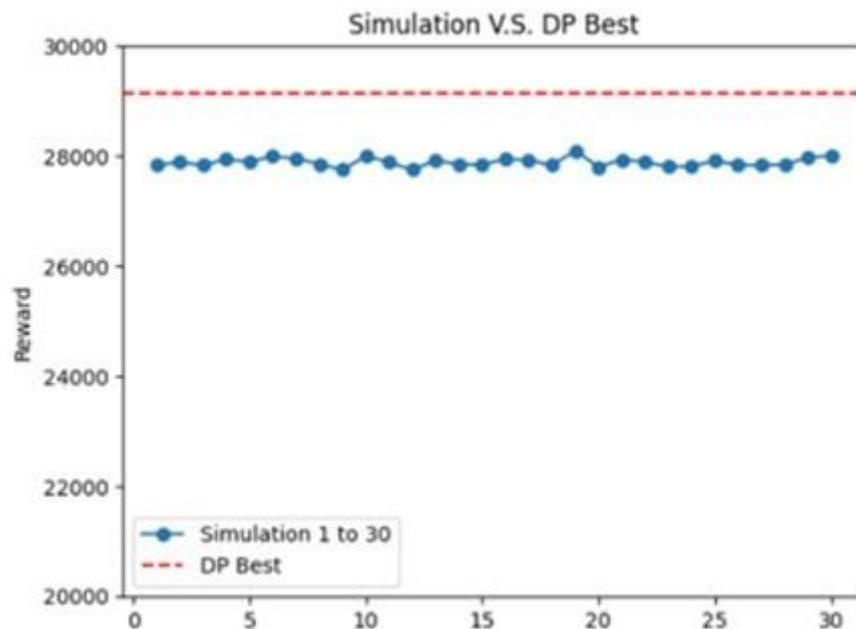


圖 20 模擬標準差為 20 之平均為 70 的 30 次模擬利潤

最後我們在已知環境參數的情況下做模擬訂價，並與動態規劃求解答案計算偏好學習成本 Gap，即為圖 21 之灰點，得到 Gap 為-0.17%，說明用已知的 μ 和 σ 做模擬訂價得到的利潤，與真實利潤差異甚小，證明整體模擬驗證流程及方法之



正確性。圖 21 之藍點則為標準差為 20 之最終模擬驗證結果，Average Gap 為-3.93%，說明本研究之方法僅需要少許的成本，即可學習真實的利潤。

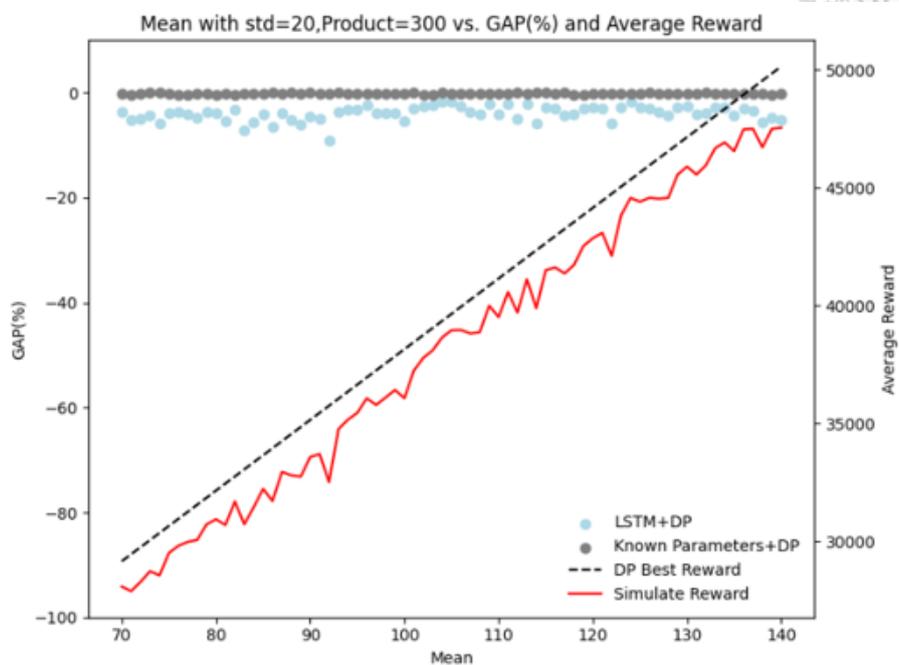


圖 21 模擬標準差為 20 之偏好學習成本圖

➤ 標準差為 30 之模擬驗證結果

將估計之 $\hat{\mu}$ 和 $\hat{\sigma}$ 計算常態分布 KL Divergence，得到結果為 0.0238，符合訓練 LSTM 模型時模型收斂之對應 Loss 值，並在當期銷售期數為 1、3、5、10、25、50、100、150 與 250 期時 LSTM 模型估計 $\hat{\mu}$ 和 $\hat{\sigma}$ 的結果做成圖 22，證明 LSTM 模型之估計效力良好。

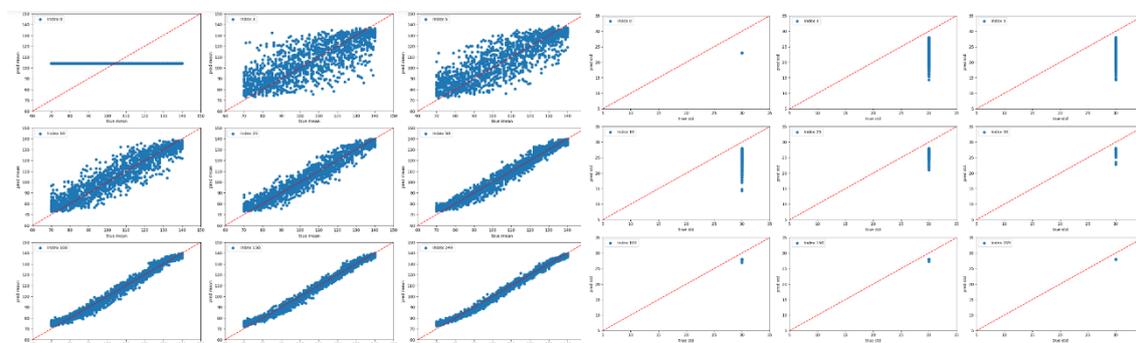


圖 22 模擬標準差為 30 之估計 250 期銷售時程的 $\hat{\mu}$ 和 $\hat{\sigma}$



接者在商品銷售模擬中，消費者願付價格 p_t^C 會服從 $N(70,30)$ ，因此我們將 30 次模擬中的第一次模擬，過程中所有產生的消費者願付價格 p_t^C 轉換成機率密度函數柱狀圖，證明 p_t^C 之正確性與真實分布相符。

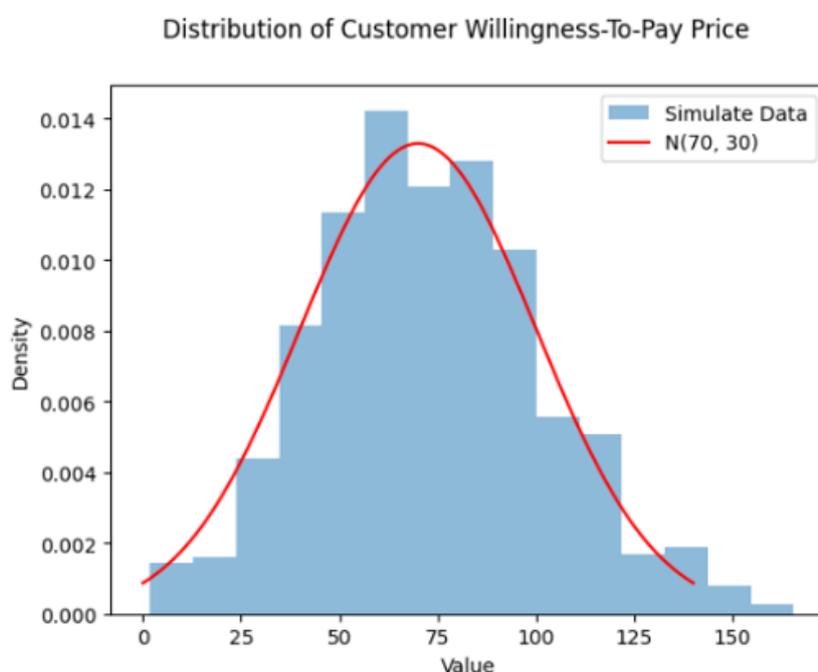


圖 23 模擬標準差為 30 之第一次模擬之消費者願付價格機率密度函數柱狀圖

再者我們將全部 71 種平均值模擬 30 次之利潤做分析。如圖 24 所示，可以得知隨著平均值增加，模擬利潤以及真實利潤皆穩定上升。由圖 25 可以得知相同平均值情況下，30 次模擬之利潤差異不大，證明本研究模擬之可靠性及正確性。

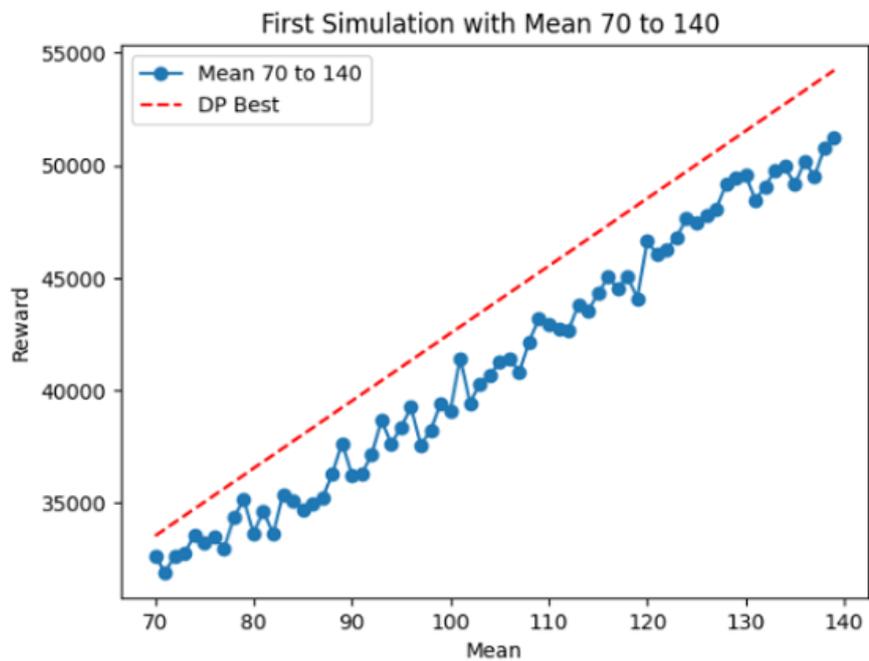
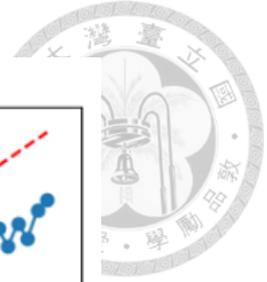


圖 24 模擬標準差為 30 之第一次模擬 71 種平均之模擬利潤

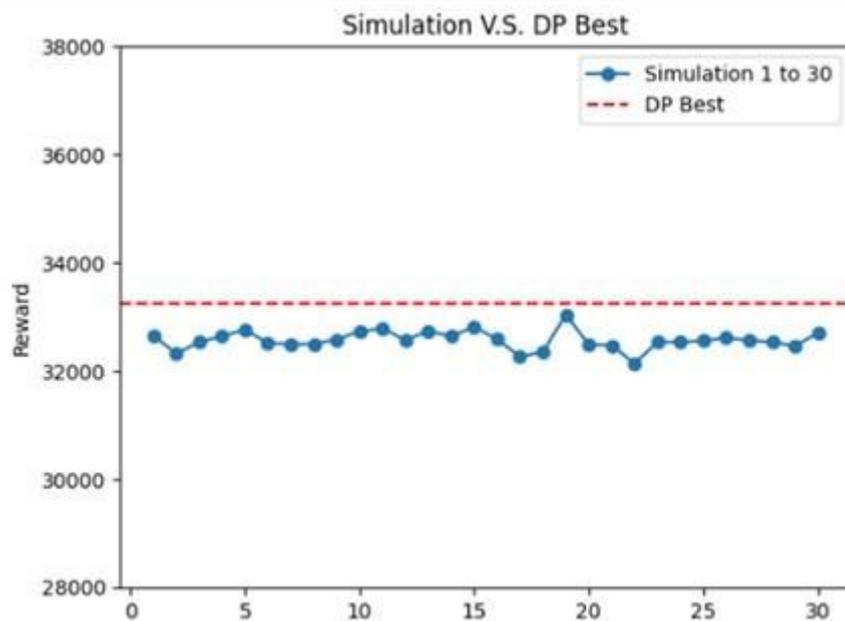


圖 25 模擬標準差為 30 之平均為 70 的 30 次模擬利潤

最後我們在已知環境參數的情況下做模擬訂價，並與動態規劃求解答案計算偏好學習成本 Gap，即為圖 26 之灰點，得到 Gap 為-0.16%，說明用已知的 μ 和 σ 做模擬訂價得到的利潤，與真實利潤差異甚小，證明整體模擬驗證流程及方法之

正確性。圖 26 之藍點則為標準差為 30 之最終模擬驗證結果，Average Gap 為-3.63%，說明本研究之方法僅需要少許的成本，即可學習真實的利潤。

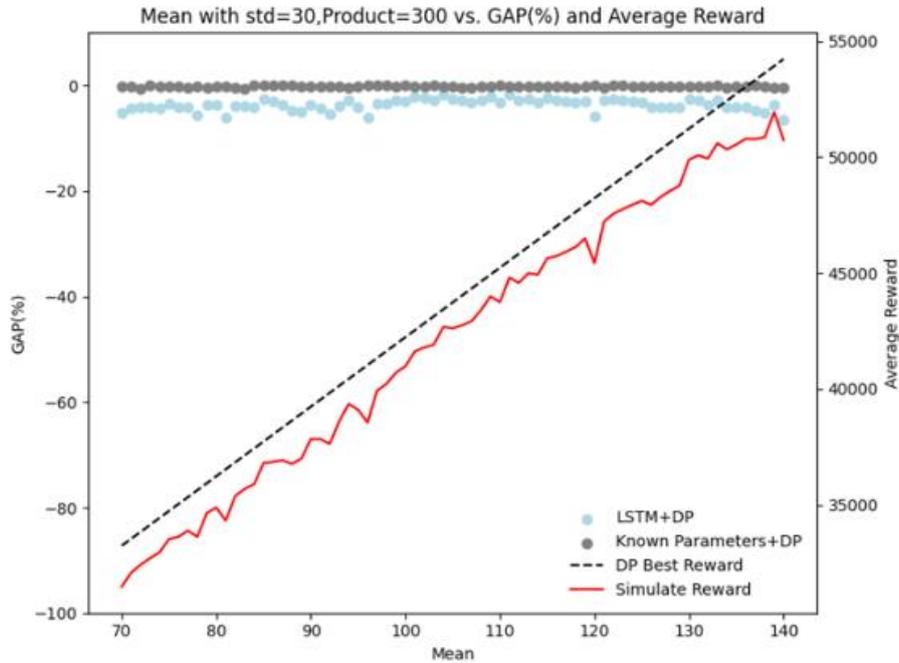


圖 26 模擬標準差為 30 之偏好學習成本圖

從以上模擬結果圖 16、圖 21、圖 26 中可以得知：

- i. 由 5.1.2 中提及 LSTM 初期估計平均值就會落在 105 附近，因此當模擬平均值範圍在 105 附近時，偏好學習成本較小。
- ii. 5.1.2 中提及 LSTM 初期估計標準差值會落在 24 附近，因此模擬標準差為 10 之偏好學習成本會較模擬標準差為 20 和 30 來的大。
- iii. 由以上模擬驗證結果得知，面對不同參數之常態分布，我們的模型求解得到之模擬利潤與真實利潤差異不到 4%。



5.2 WTP 服從伽瑪分布

5.2.1 參數設置

本小節將會對於 3.2.1 中 LSTM 消費者購買資料模擬的參數設置進行介紹，其中消費者支付意願分布服從伽瑪分布，參數之設置如表 5 所示。

Parameters	Value
Planning Horizon	250
Number of Products	300
Arrival Probability	0.7
Shape of WTP	10~30
Scale of WTP	2~5
Range of Price	10~200
Price Difference	1
Maximum Number of Customers	20

表 5 WTP 為伽瑪分布之參數設置

消費者支付意願服從伽瑪分布之參數設置與常態分布大致相同，不同的地方是商品售價下界調整為 10，售價上界仍為 200。其中商品售價也是依照伽瑪分布之參數範圍設置，由於伽瑪分布的平均值是由兩參數相乘決定的，因此將商品售價下界調整為 10，使模型有足夠多的有效資料做學習。



5.2.2 長短期記憶神經網路模型測試

圖 27 及圖 28 為當期銷售期數為 1、3、5、10、25、50、100、150 與 250 期時 LSTM 模型估計的結果，其中橫軸為估計環境形狀參數值和尺度參數值，縱軸為真實環境形狀參數值和尺度參數值，紅色斜虛線為標準答案，每一個藍色點代表其中一筆測試資料在當期時 LSTM 模型估計之值，越靠近紅色斜虛線代表估計越準確。其中伽瑪分布的特性為其平均值是由形狀參數乘上尺度參數得到，因此多增加一張形狀參數乘上尺度參數之模型估計圖，驗證其準確性。由圖 29 可以觀察到在完全沒有參考數據的第 1 期，估計會落在兩參數相乘範圍的中心位置附近，並且在 50 期時估計就已經非常接近真實值，而第 250 期可以看到估計十分準確。

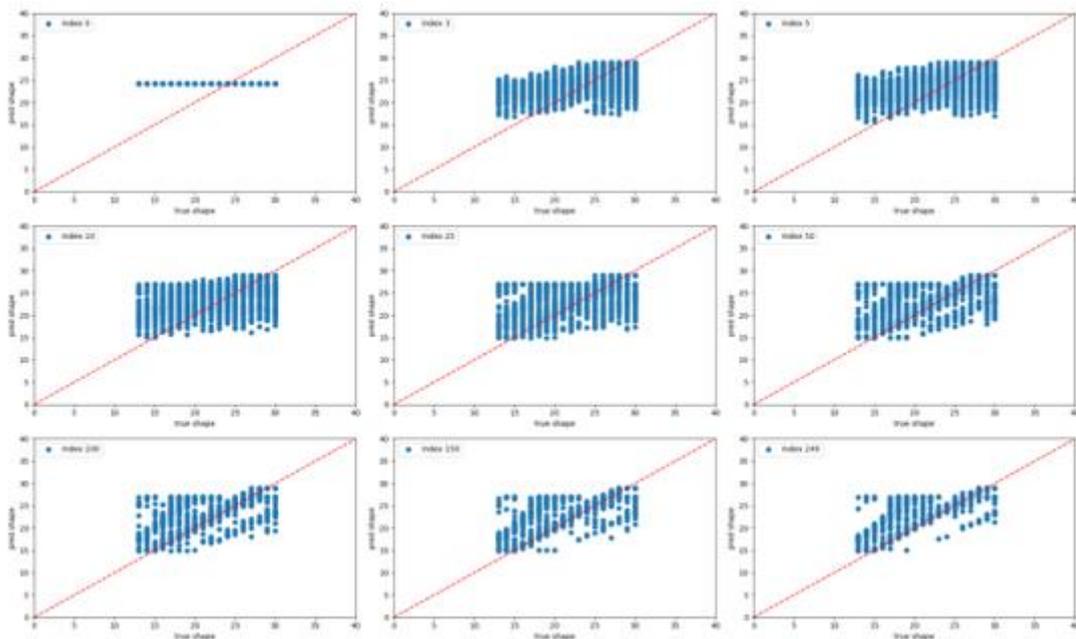


圖 27 估計 250 期銷售時程之伽瑪分布形狀參數值

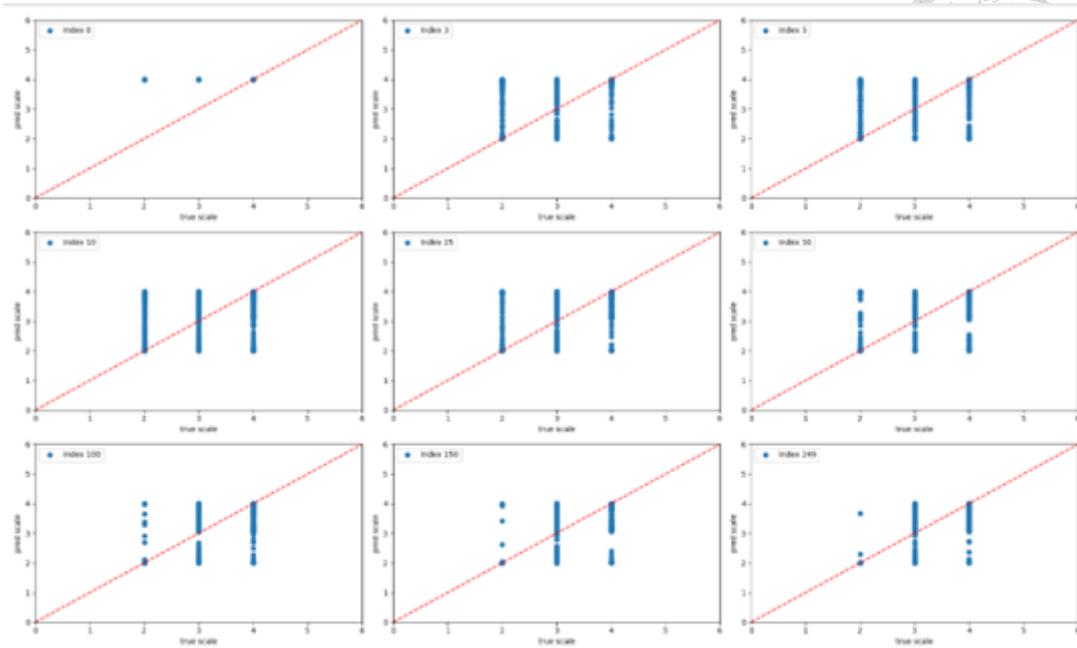


圖 28 估計 250 期銷售時程之伽瑪分布尺度參數值

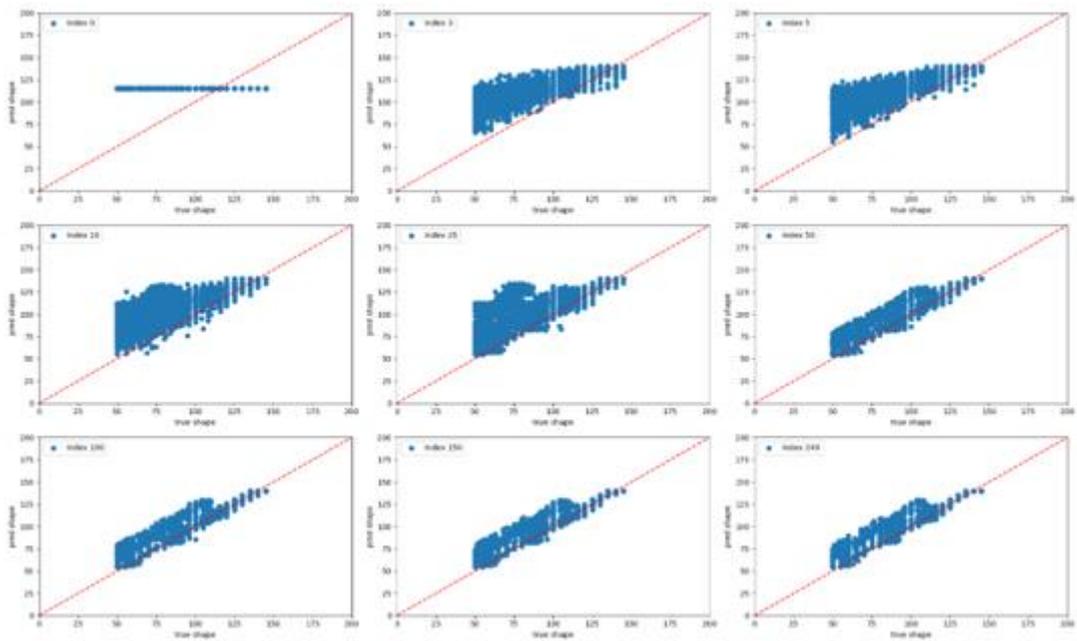


圖 29 估計 250 期銷售時程之伽瑪分布形狀參數乘上尺度參數之值



5.2.3 模擬驗證流程

本研究將使用離散事件模擬，模擬形狀參數範圍 10-30，尺度參數 2-5 總共 84 種環境。每一種環境皆模擬 30 次，以下將進行模擬流程之說明，其中模擬過程中會用到之參數如表 4 所示。

- i. 在未知的環境市場中使用 LSTM 模型，進行環境參數識別，得到估計之伽瑪分布參數 $\hat{\alpha}$ 和 $\hat{\theta}$ ，並使用估計之環境參數得到當期售價 p_t 。
- ii. 使用已知環境參數 α 和 θ 、當期售價 p_t 、消費者到達率 λ 、最大來客數 C^M 以及目前剩餘存貨 m_t 進行商品銷售模擬。
- iii. 每一單位時間若有消費者到達，此消費者的消費者願付價格 p_t^C 服從伽瑪分布。
- iv. 當消費者的願付價格 $p_t^C \geq$ 當期售價 p_t 時，此消費者購買一個商品並離開，否則此消費者會直接離開，且每位消費者最多購買一個商品。
- v. 當商品銷售完或是售出數量 $n_t <$ 存貨數量 m_t 時，則視為購買結束，此時當期模擬利潤 $r_t =$ 商品售出數量 $n_t * 當期售價 p_t 。$
- vi. 紀錄當期售價 p_t 、消費者到達數量 l_t 、商品售出數量 n_t 作為下一期之 LSTM Input。
- vii. $\hat{Q} = \sum r_t$ 。
- viii. Q 為每次模擬時，用已知環境參數 α 和 θ 做動態規劃訂價求解得到之最佳利潤。
- ix. $R = \frac{\sum \hat{B}}{\text{模擬次數 } 30}$ 。
- x. $\hat{R} = \frac{\sum B}{\text{模擬次數 } 30}$ 。
- xi. 計算 Gap，得到學習真實利潤需要的成本。



5.2.4 模擬驗證結果

本研究伽瑪分布則是模擬形狀參數範圍 10-30，尺度參數為 2-5。本小節將詳細介紹模擬驗證結果，並使用模擬過程中的各種輔助圖，證明偏好學習成本 Gap 結果無誤。

➤ 尺度參數為 2 之模擬驗證結果

如 5.2.3 提及之模擬流程，首先會使用 LSTM 模型估計環境參數，得到估計之 $\hat{\alpha}$ 和 $\hat{\theta}$ ，因此將估計之 $\hat{\alpha}$ 和 $\hat{\theta}$ 計算常態分布 KL Divergence，得到結果為 11.1027，符合訓練 LSTM 模型時模型收斂之對應 Loss 值，並在當期銷售期數為 1、3、5、10、25、50、100、150 與 250 期時 LSTM 模型估計 $\hat{\alpha}$ 和 $\hat{\theta}$ 的結果做成圖 31，以及將 $\hat{\alpha} * \hat{\theta}$ 做成圖 32，證明 LSTM 模型之估計效力良好。

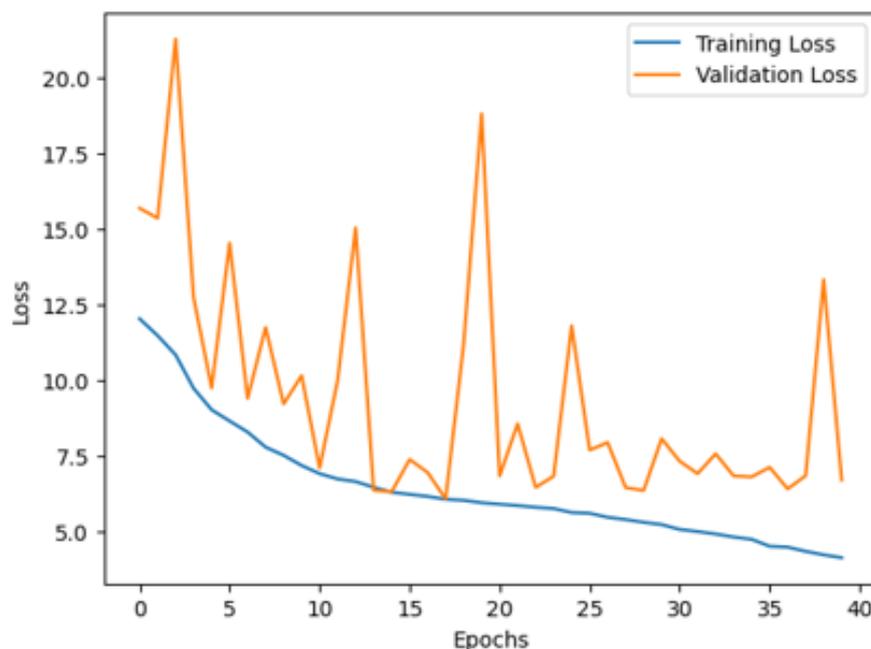


圖 30 LSTM 模型之訓練伽瑪分布 Loss 圖

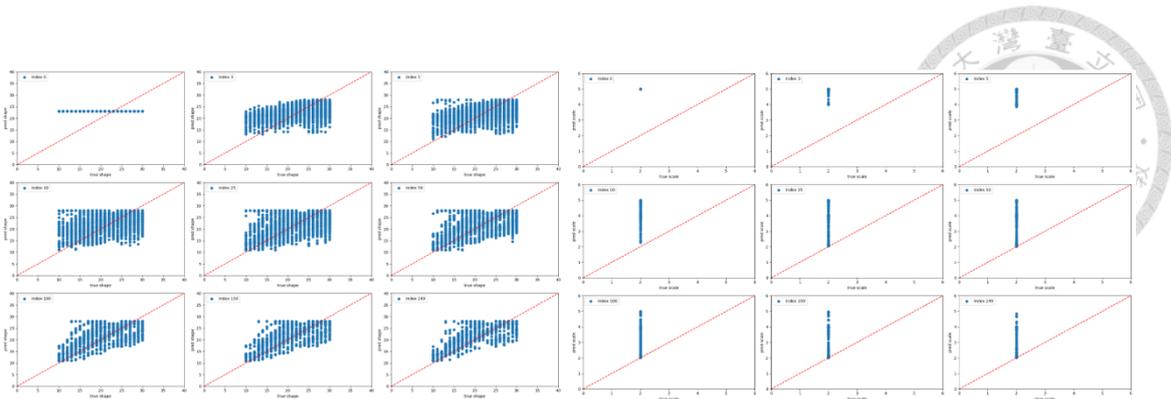


圖 31 模擬尺度參數為 2 之估計 250 期銷售時程的 $\hat{\alpha}$ 和 $\hat{\theta}$

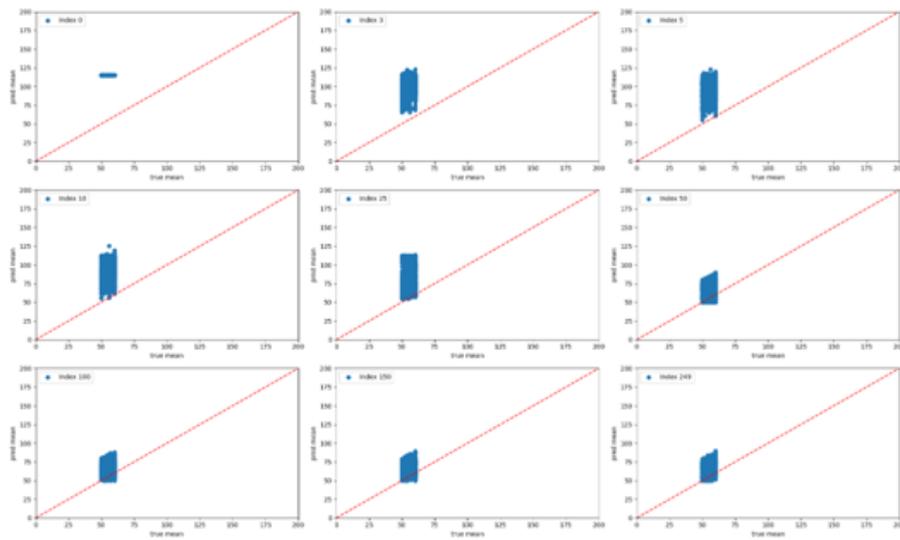


圖 32 模擬尺度參數為 2 之估計 250 期銷售時程的 $\hat{\alpha} * \hat{\theta}$

接者在商品銷售模擬中，消費者願付價格 p_t^C 會服從 $\Gamma(10,2)$ ，因此我們將 30 次模擬中的第一次模擬，過程中所有產生的消費者願付價格 p_t^C 轉換成機率密度函數柱狀圖，證明 p_t^C 之正確性與真實分布相符。

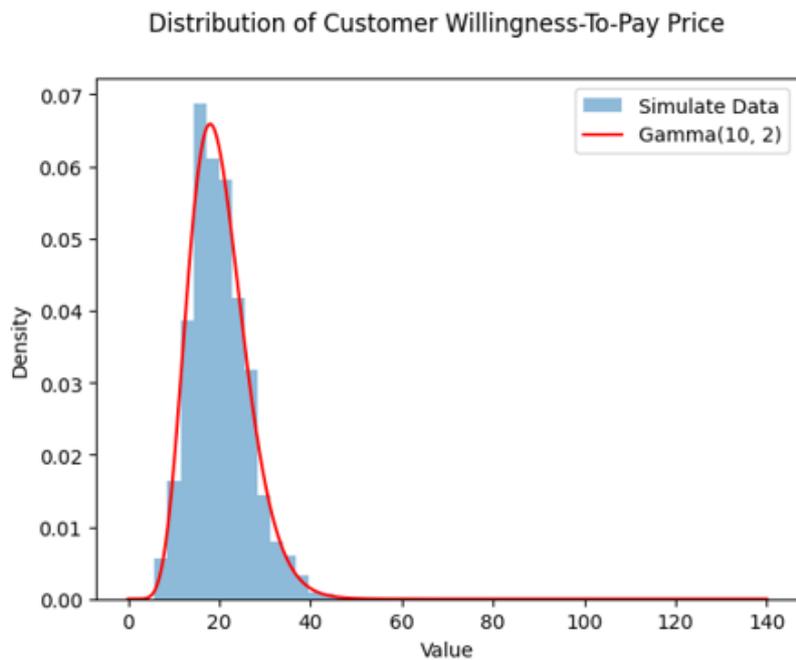


圖 33 模擬尺度參數為 2 之第一次模擬之消費者願付價格機率密度函數柱狀圖

再者我們將全部 21 種平均值模擬 30 次之利潤做分析。如圖 34 所示，藍點為第一次模擬，平均 10-30 總共 21 種模擬利潤，紅色虛線為動態規劃求解之真實利潤，可以得知隨著平均值增加，模擬利潤以及真實利潤皆穩定上升，而形狀參數為 29 和 30 時會造成伽瑪分布稍微超出商品售價上界，因此模擬利潤與真實利潤差異較大。圖 35 藍點為形狀參數為 10 之 30 次模擬利潤，紅色虛線為動態規劃求解之真實利潤，可以得知相同形狀參數情況下，30 次模擬之利潤差異不大，證明本研究模擬之可靠性及正確性。

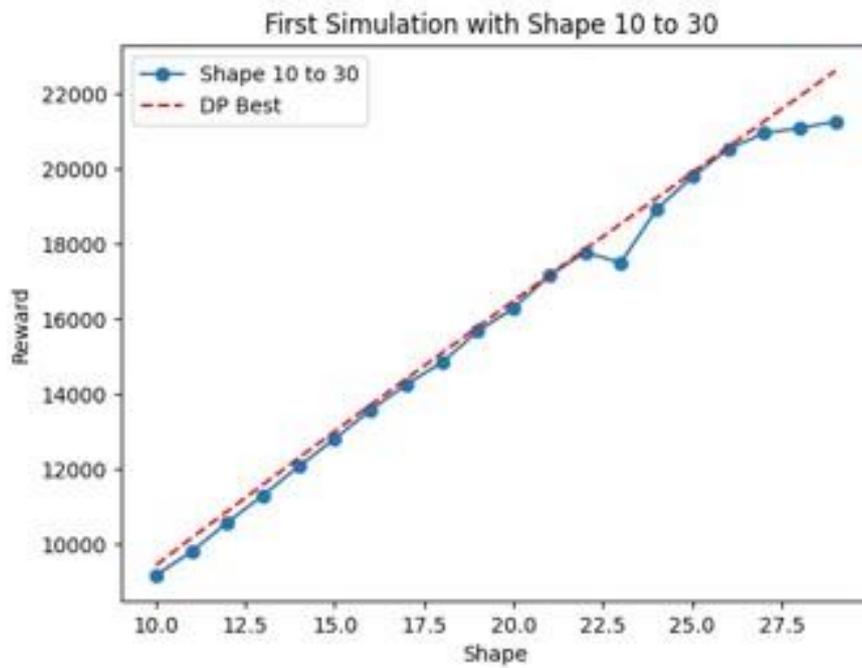
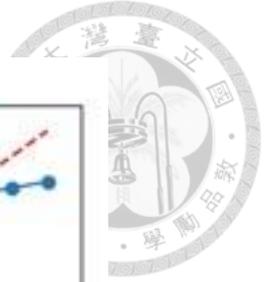


圖 34 模擬尺度參數為 2 之第一次模擬 21 種形狀參數之模擬利潤

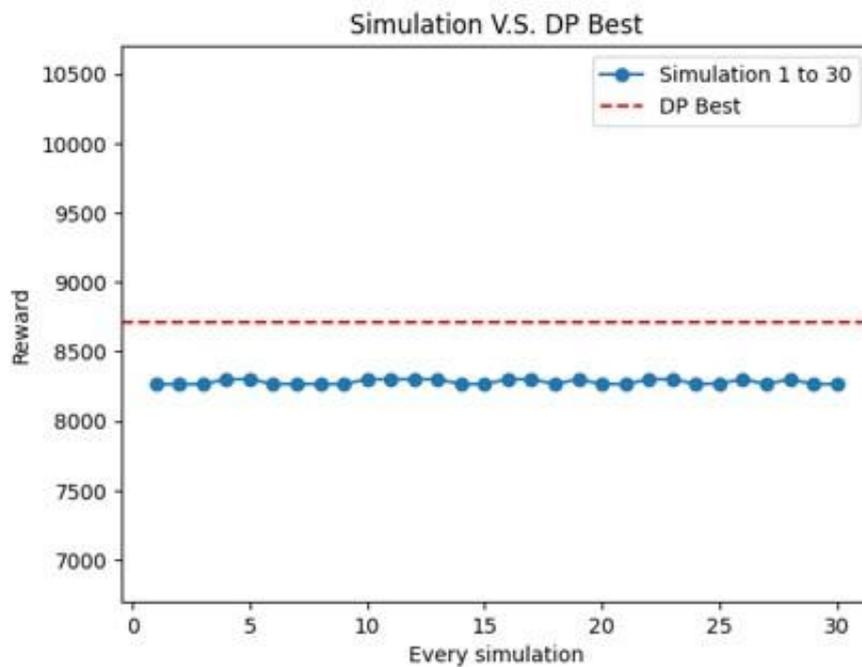


圖 35 模擬尺度參數為 2 之形狀參數為 10 的 30 次模擬利潤

最後我們在已知環境參數的情況下做模擬訂價，並與動態規劃求解答案計算偏好學習成本 Gap，即為圖 36 之灰點，得到 Gap 為-0.3%，說明用已知的 α 和 θ 做模擬訂價得到的利潤，與真實利潤差異甚小，證明整體模擬驗證流程及方法之正



確性。圖 36 之藍點則為尺度參數為 2 之最終模擬驗證結果，Average Gap 為-4.87%，說明本研究之方法僅需要少許的成本，即可學習真實的利潤。另外黑色虛線是動態規劃求解之真實利潤，紅色實線是模擬 30 次之平均利潤。

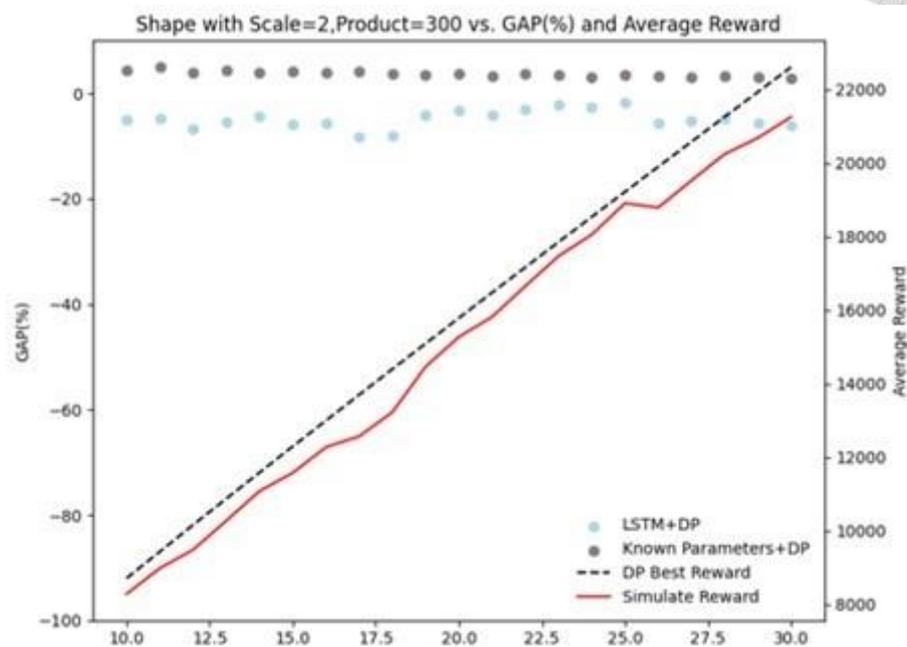


圖 36 模擬尺度參數為 2 之偏好學習成本圖

➤ 尺度參數為 3 之模擬驗證結果

將估計之 $\hat{\alpha}$ 和 $\hat{\theta}$ 計算常態分布 KL Divergence，得到結果為 10.0732，符合訓練 LSTM 模型時模型收斂之對應 Loss 值，並在當期銷售期數為 1、3、5、10、25、50、100、150 與 250 期時 LSTM 模型估計 $\hat{\alpha}$ 和 $\hat{\theta}$ 的結果做成圖 37，以及將 $\hat{\alpha}*\hat{\theta}$ 做成圖 38，證明 LSTM 模型之估計效力良好。證明 LSTM 模型之估計效力良好。

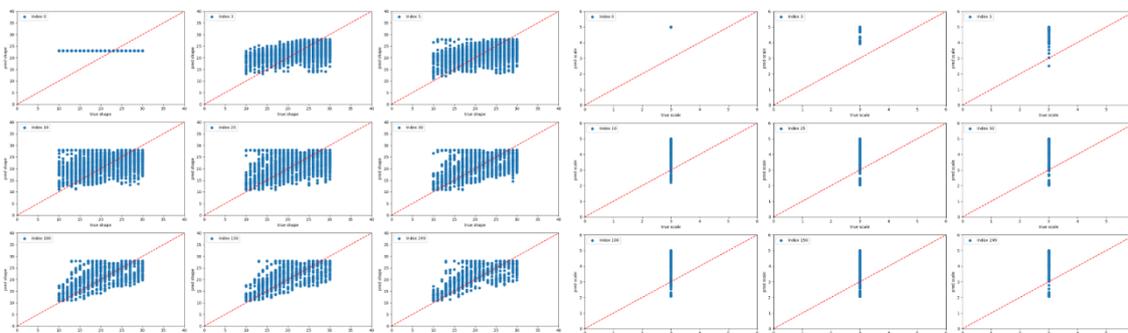


圖 37 模擬尺度參數為 3 之估計 250 期銷售時程的 $\hat{\alpha}$ 和 $\hat{\theta}$

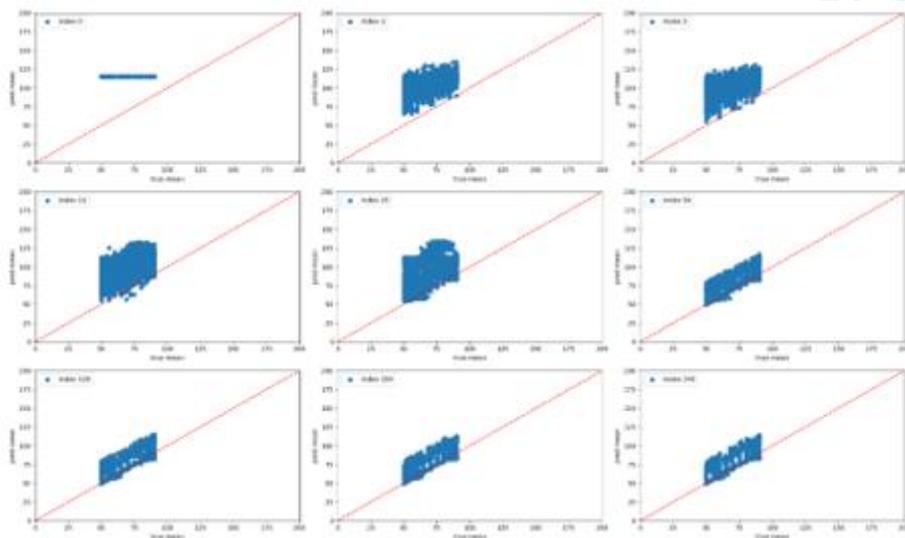
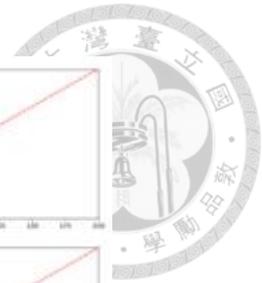


圖 38 模擬尺度參數為 3 之估計 250 期銷售時程的 $\hat{\alpha} * \hat{\theta}$

接者在商品銷售模擬中，消費者願付價格 p_t^C 會服從 $\Gamma(10,3)$ ，因此我們將 30 次模擬中的第一次模擬，過程中所有產生的消費者願付價格 p_t^C 轉換成機率密度函數柱狀圖，證明 p_t^C 之正確性與真實分布相符。

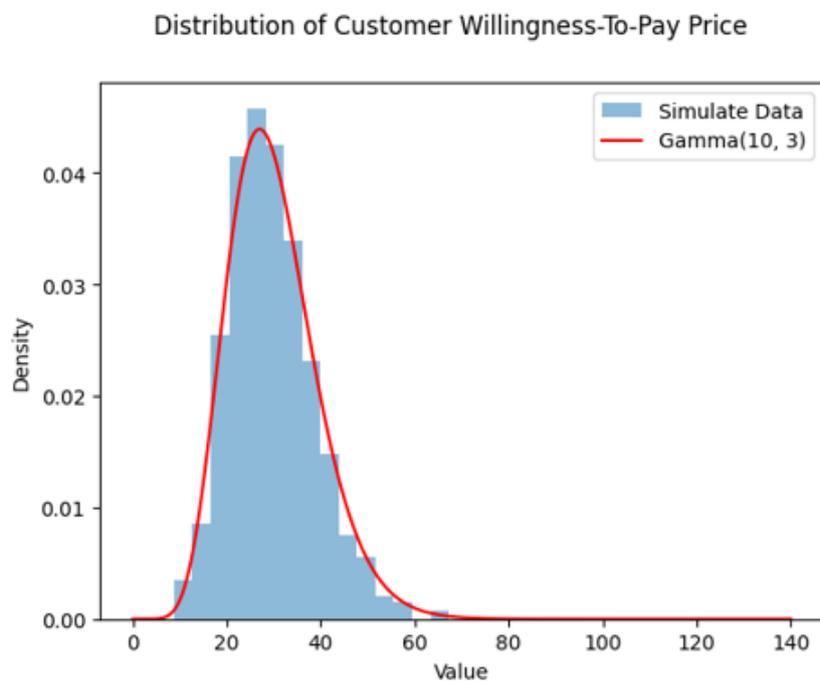


圖 39 模擬尺度參數為 3 之第一次模擬之消費者願付價格機率密度函數柱狀圖

再者我們將全部 21 種形狀參數模擬 30 次之利潤做分析。如圖 40 所示，可以得知隨著形狀參數增加，模擬利潤以及真實利潤皆穩定上升，而形狀參數為 29



和 30 時會造成伽瑪分布稍微超出商品售價上界，因此模擬利潤與真實利潤差異較大。由圖 41 可以得知相同尺度參數情況下，30 次模擬之利潤差異不大，證明本研究模擬之可靠性及正確性。

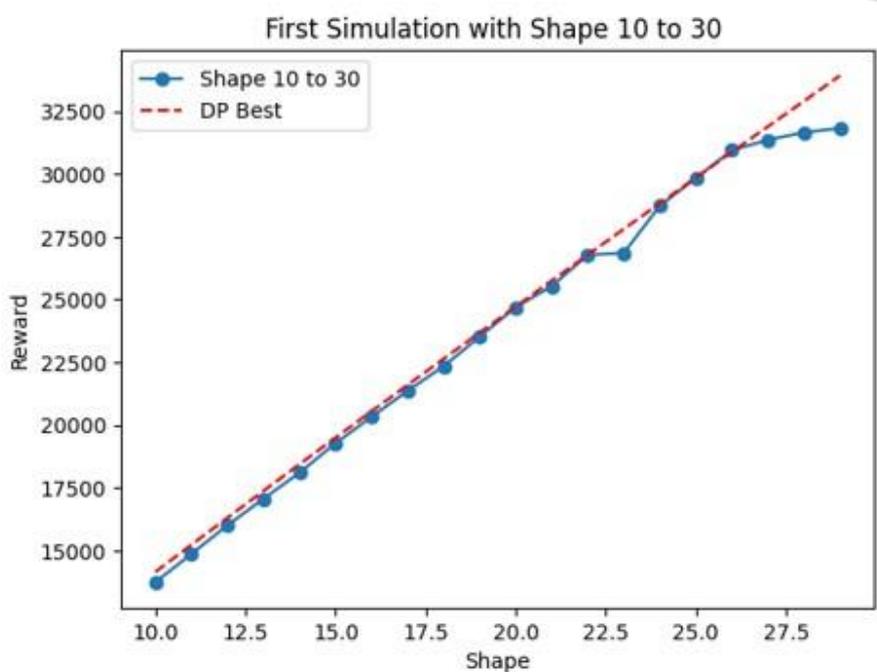


圖 40 模擬尺度參數為 3 之第一次模擬 21 種形狀參數之模擬利潤

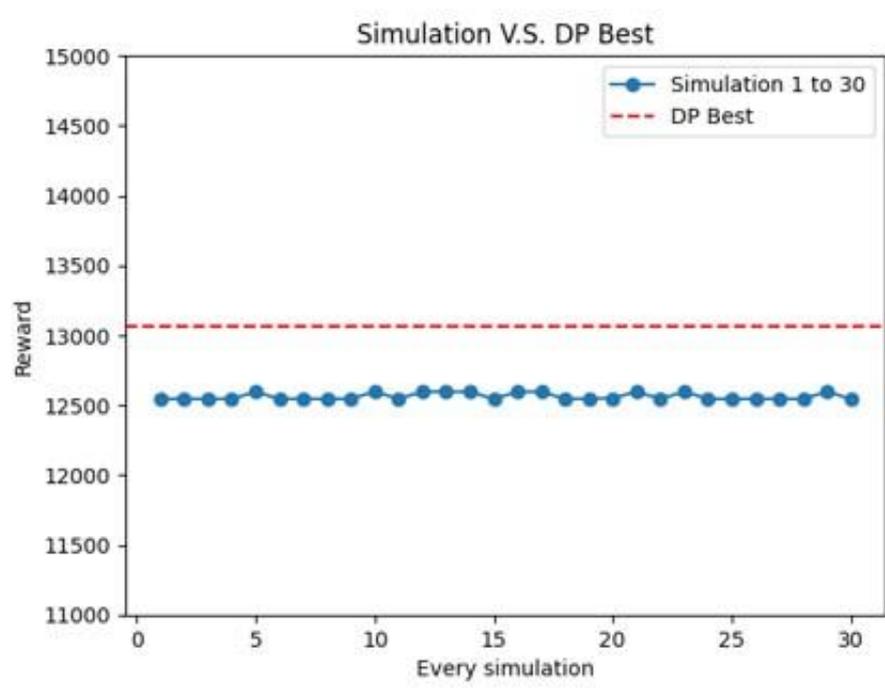


圖 41 模擬尺度參數為 3 之形狀參數為 10 的 30 次模擬利潤

最後我們在已知環境參數的情況下做模擬訂價，並與動態規劃求解答案計算偏好學習成本 Gap，即為圖 42 之灰點，得到 Gap 為-0.16%，說明用已知的 α 和 θ 做模擬訂價得到的利潤，與真實利潤差異甚小，證明整體模擬驗證流程及方法之正確性。圖 42 之藍點則為尺度參數為 3 之最終模擬驗證結果，Average Gap 為-4.15%，說明本研究之方法僅需要少許的成本，即可學習真實的利潤。

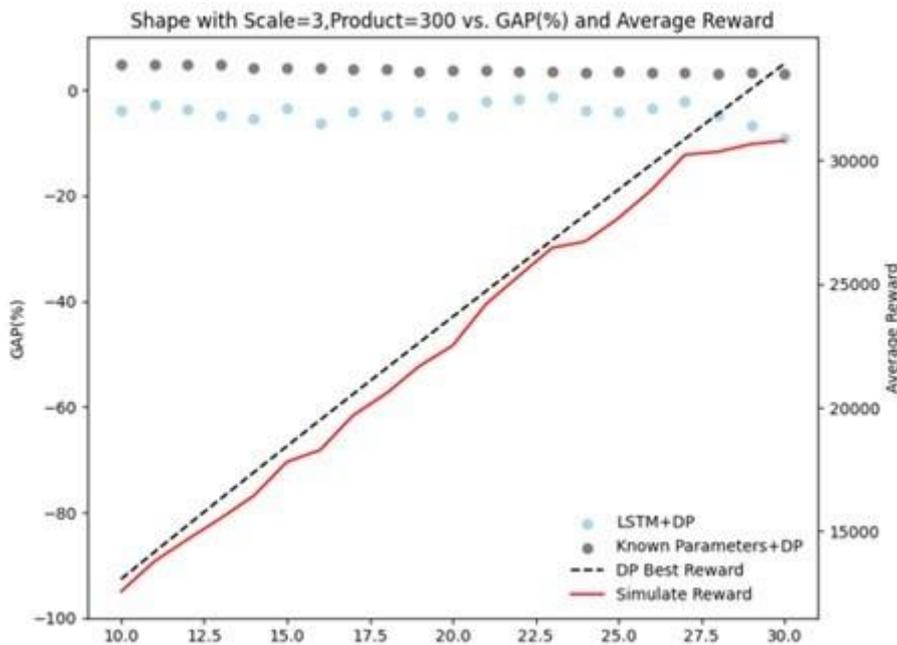


圖 42 模擬尺度參數為 3 之偏好學習成本圖

➤ 尺度參數為 4 之模擬驗證結果

將估計之 $\hat{\alpha}$ 和 $\hat{\theta}$ 計算常態分布 KL Divergence，得到結果為 10.3821，符合訓練 LSTM 模型時模型收斂之對應 Loss 值，並在當期銷售期數為 1、3、5、10、25、50、100、150 與 250 期時 LSTM 模型估計 $\hat{\alpha}$ 和 $\hat{\theta}$ 的結果做成圖 43，以及將 $\hat{\alpha}*\hat{\theta}$ 做成圖 44，證明 LSTM 模型之估計效力良好。證明 LSTM 模型之估計效力良好。

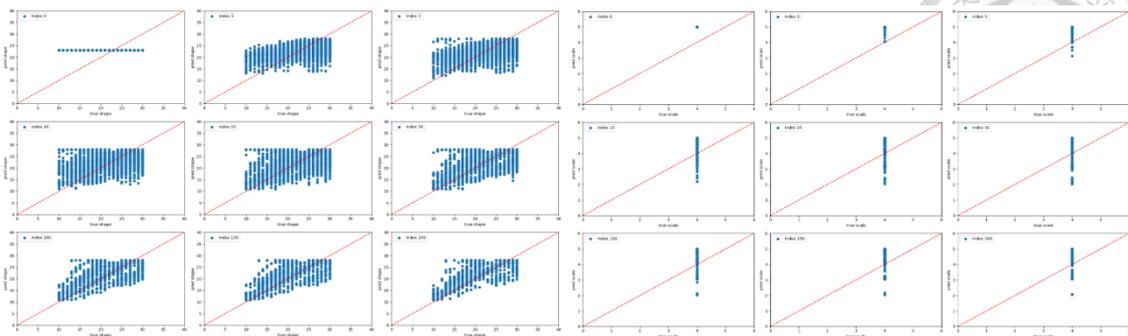


圖 43 模擬尺度參數為 4 之估計 250 期銷售時程的 $\hat{\alpha}$ 和 $\hat{\theta}$

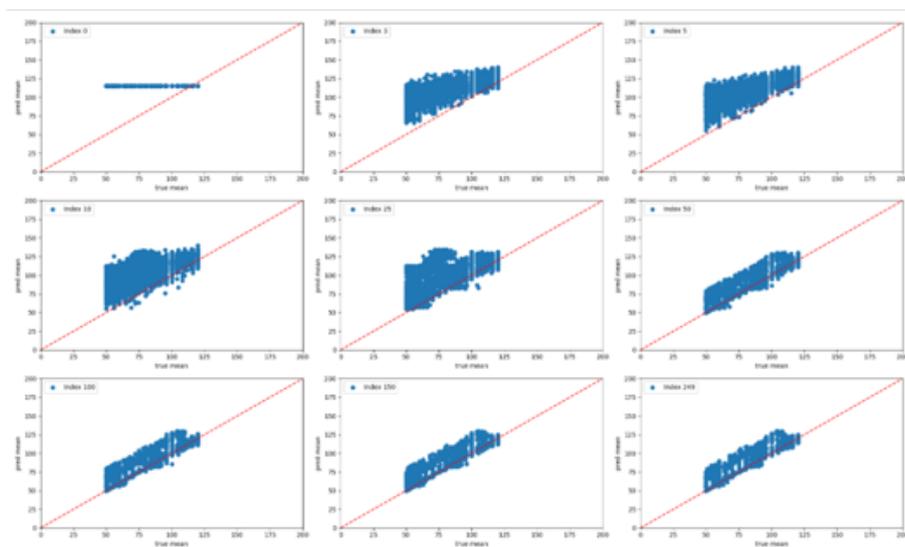


圖 44 模擬尺度參數為 4 之估計 250 期銷售時程的 $\hat{\alpha} * \hat{\theta}$

接者在商品銷售模擬中，消費者願付價格 p_t^C 會服從 $\Gamma(10,4)$ ，因此我們將 30 次模擬中的第一次模擬，過程中所有產生的消費者願付價格 p_t^C 轉換成機率密度函數柱狀圖，證明 p_t^C 之正確性與真實分布相符。



圖 45 模擬尺度參數為 4 之第一次模擬之消費者願付價格機率密度函數柱狀圖

再者我們將全部 21 種形狀參數模擬 30 次之利潤做分析。如圖 46 所示，可以得知隨著形狀參數增加，模擬利潤以及真實利潤皆穩定上升，而形狀參數為 29 和 30 時會造成伽瑪分布稍微超出商品售價上界，因此模擬利潤與真實利潤差異較大。由圖 47 可以得知相同尺度參數情況下，30 次模擬之利潤差異不大，證明本研究模擬之可靠性及正確性。

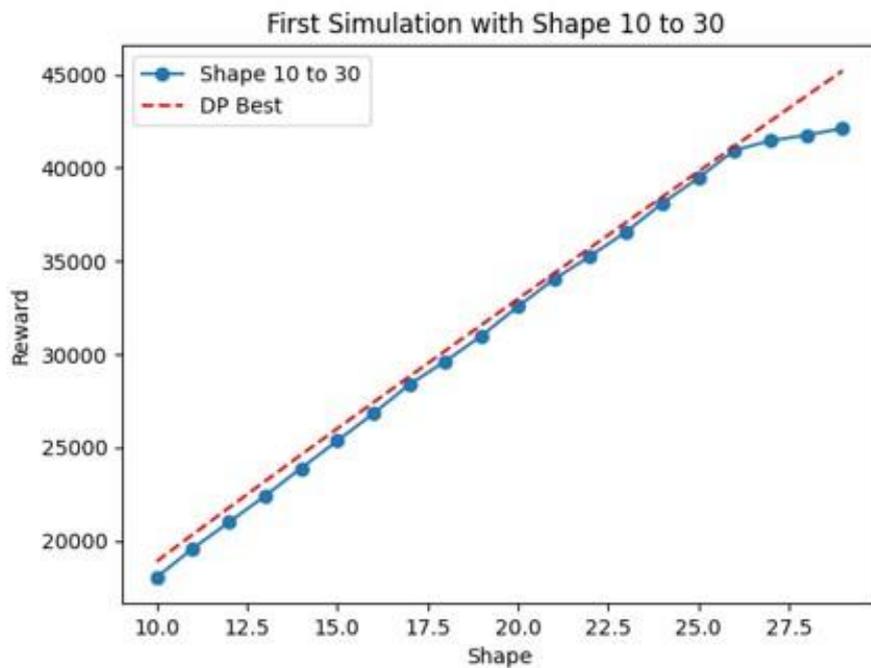
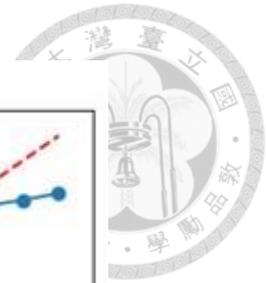


圖 46 模擬尺度參數為 4 之第一次模擬 21 種形狀參數之模擬利潤

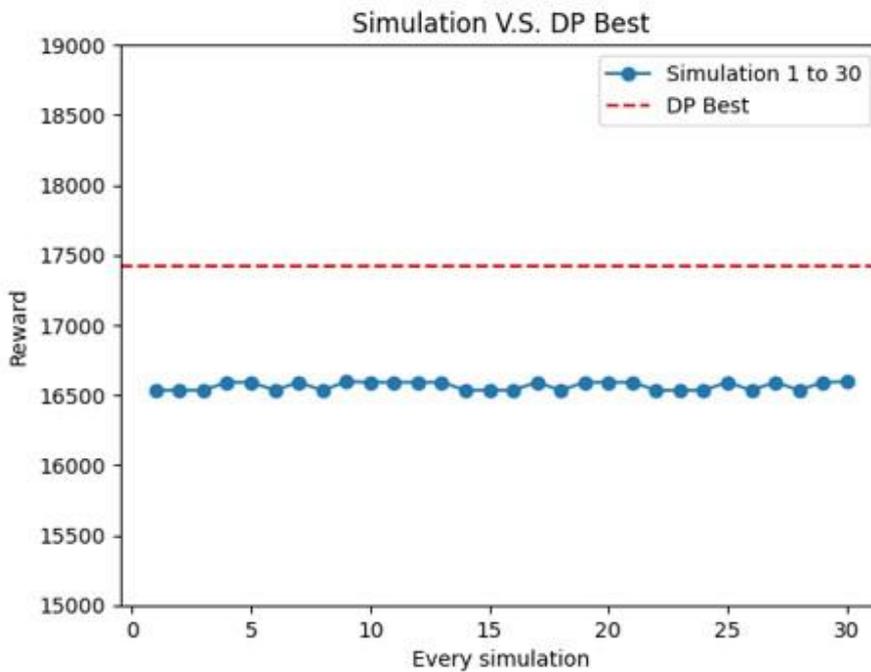


圖 47 模擬尺度參數為 4 之形狀參數為 10 的 30 次模擬利潤

最後我們在已知環境參數的情況下做模擬訂價，並與動態規劃求解答案計算偏好學習成本 Gap，即為圖 48 之灰點，得到 Gap 為-0.16%，說明用已知的 α 和 θ 做模擬訂價得到的利潤，與真實利潤差異甚小，證明整體模擬驗證流程及方法之

正確性。圖 48 之藍點則為尺度參數為 4 之最終模擬驗證結果，Average Gap 為-4.6%，說明本研究之方法僅需要少許的成本，即可學習真實的利潤。

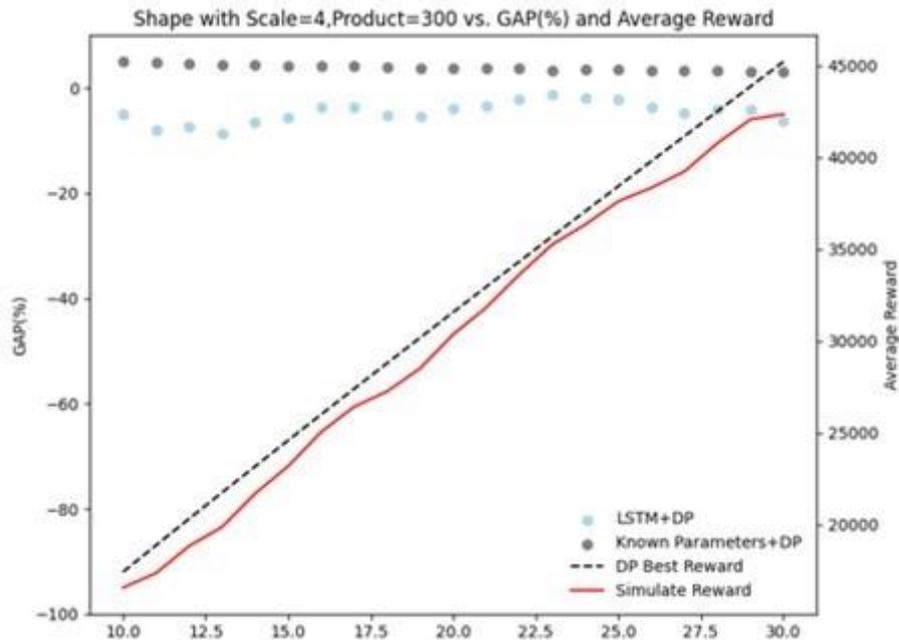


圖 48 模擬尺度參數為 4 之偏好學習成本圖

➤ 尺度參數為 5 之模擬驗證結果

將估計之 $\hat{\alpha}$ 和 $\hat{\theta}$ 計算常態分布 KL Divergence，得到結果為 9.7143，符合訓練 LSTM 模型時模型收斂之對應 Loss 值，並在當期銷售期數為 1、3、5、10、25、50、100、150 與 250 期時 LSTM 模型估計 $\hat{\alpha}$ 和 $\hat{\theta}$ 的結果做成圖 49，以及將 $\hat{\alpha}*\hat{\theta}$ 做成圖 50，證明 LSTM 模型之估計效力良好。證明 LSTM 模型之估計效力良好。

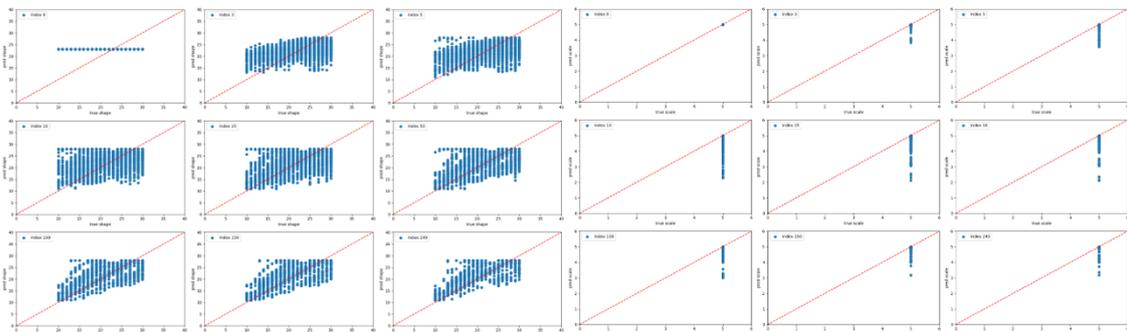


圖 49 模擬尺度參數為 5 之估計 250 期銷售時程的 $\hat{\alpha}$ 和 $\hat{\theta}$

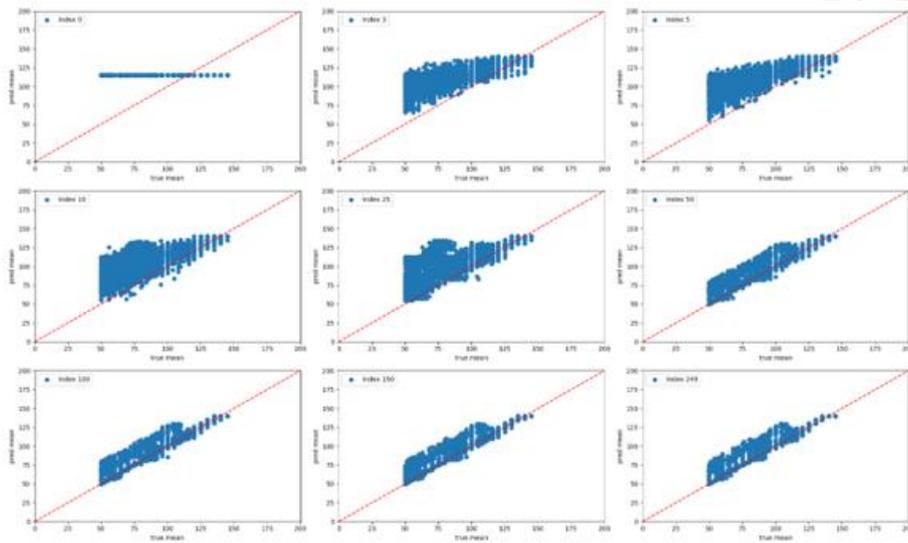
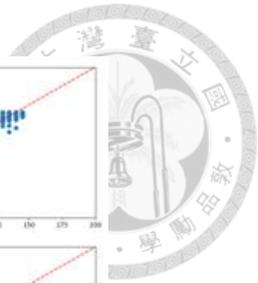


圖 50 模擬尺度參數為 5 之估計 250 期銷售時程的 $\hat{\alpha} * \hat{\theta}$

接者在商品銷售模擬中，消費者願付價格 p_t^C 會服從 $\Gamma(10,5)$ ，因此我們將 30 次模擬中的第一次模擬，過程中所有產生的消費者願付價格 p_t^C 轉換成機率密度函數柱狀圖，證明 p_t^C 之正確性與真實分布相符。

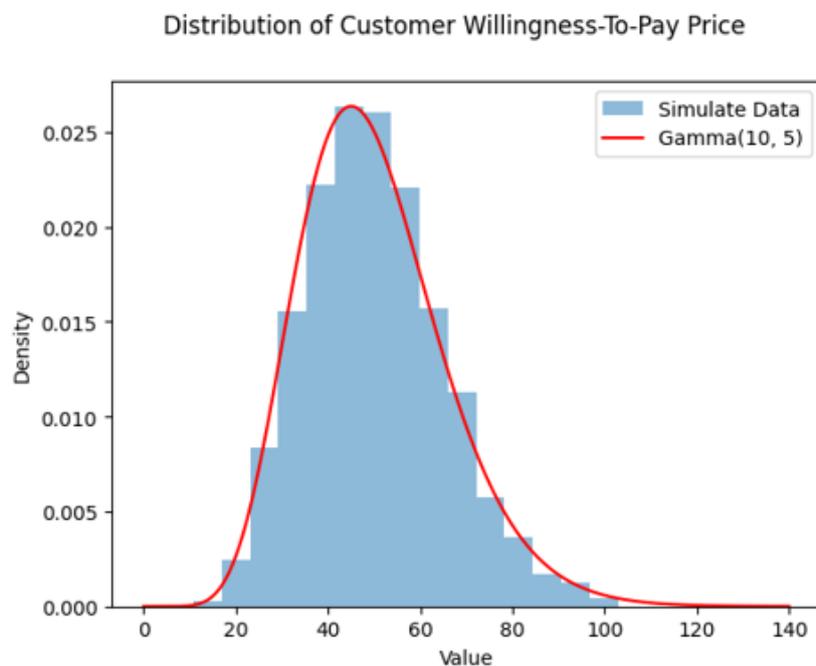


圖 51 模擬尺度參數為 5 之第一次模擬之消費者願付價格機率密度函數柱狀圖

再者我們將全部 21 種形狀參數模擬 30 次之利潤做分析。如圖 52，可以得知隨著形狀參數增加，模擬利潤以及真實利潤皆穩定上升，而形狀參數為 29 和 30



時會造成伽瑪分布稍微超出商品售價上界，因此模擬利潤與真實利潤差異較大。
由圖 53 可以得知相同尺度參數情況下，30 次模擬之利潤差異不大，證明本研究
模擬之可靠性及正確性。

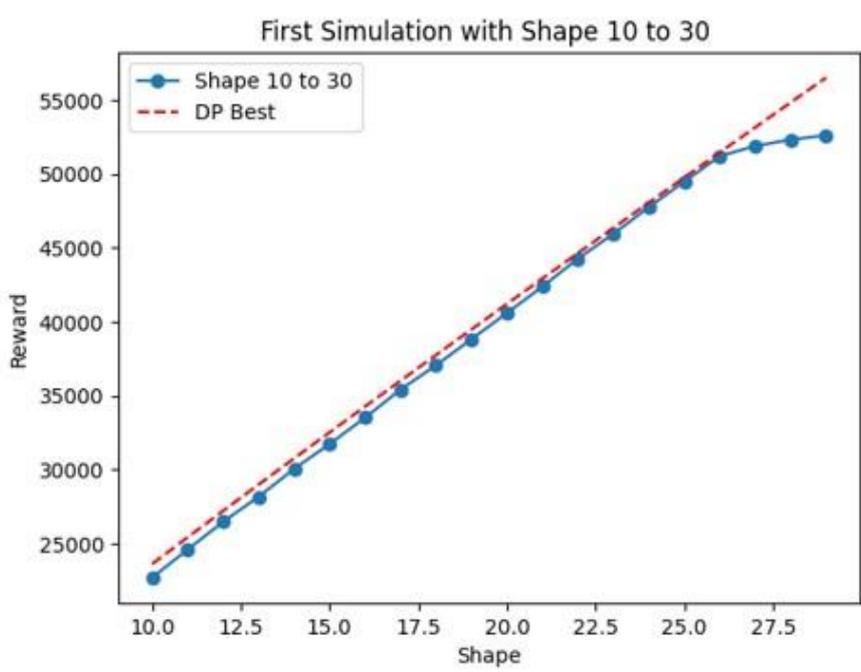


圖 52 模擬尺度參數為 5 之第一次模擬 21 種形狀參數之模擬利潤

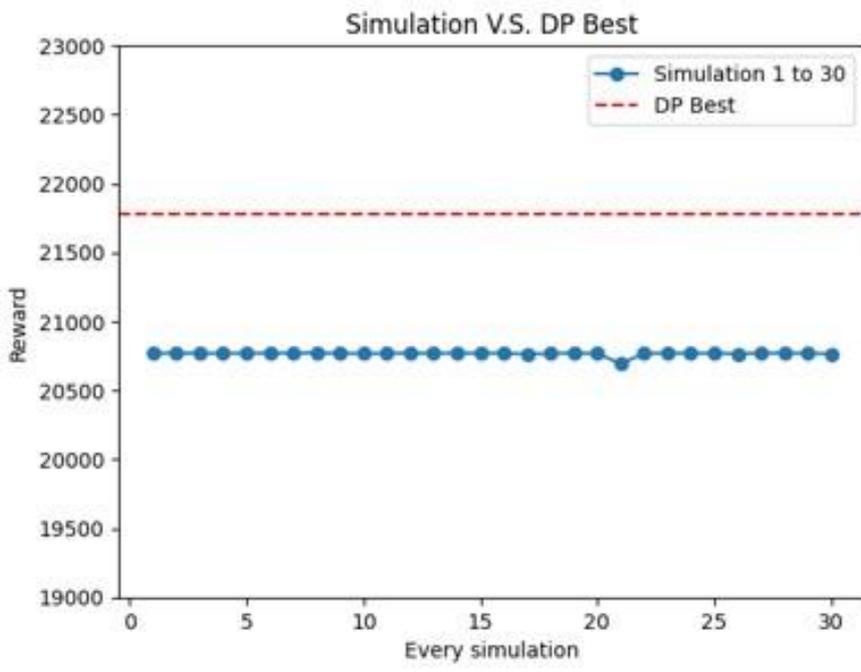


圖 53 模擬尺度參數為 5 之形狀參數為 10 的 30 次模擬利潤

最後我們在已知環境參數的情況下做模擬訂價，並與動態規劃求解答案計算偏好學習成本 Gap，即為圖 54 之灰點，得到 Gap 為-0.14%，說明用已知的 α 和 θ 做模擬訂價得到的利潤，與真實利潤差異甚小，證明整體模擬驗證流程及方法之正確性。圖 54 之藍點則為尺度參數為 5 之最終模擬驗證結果，Average Gap 為-3.75%，說明本研究之方法僅需要少許的成本，即可學習真實的利潤。

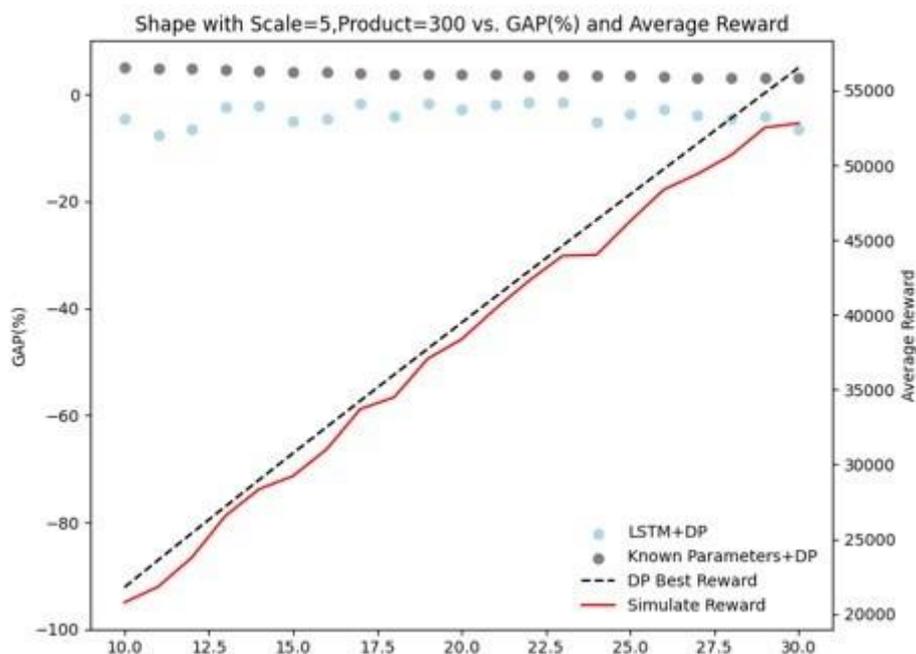


圖 54 模擬尺度參數為 5 之偏好學習成本圖

比較以上模擬結果圖 35、圖 41、圖 47 及圖 53 中可以得知:

- i. 由 5.1.2 中提及 LSTM 初期估計尺度參數值會落在 4-5 附近，因此模擬訂價時，尺度參數較大學習真實利潤所需要花費的成本較低。
- ii. 由以上模擬驗證結果得知，面對參數之伽瑪分布，我們的模型求解得到之模擬利潤與真實利潤差異不到 5%。
- iii. 由於伽瑪分布特性為其平均值是由形狀參數乘上尺度參數所得，因此當其平均值過大時(如形狀參數為 30*尺度參數為 5=平均值為在 150)會造成 WTP 分布超出商品售價上界 200，造成模擬利潤與真實利潤差異較顯著。

第六章 結論與未來方向



本研究將做出最後的總結與本研究之貢獻，並給出未來的研究方向。

6.1 結論

本研究透過模擬大量資料並萃取特徵作為訓練資料，以解決需求學習資料不足之問題。透過使用長短期記憶神經網路對於市場需求進行學習，並在未知環境中識別出消費者支付意願分布參數，作為後續動態規劃訂價之使用，並比較模擬利潤與真實利潤之間的差異。

本研究透過長短期記憶神經網路學習市場需求，並且將模型輸出為可能值之機率，再透過轉換變成期望估計值帶入 KL Divergence 做損失函數計算，使模型估計分布能越來越趨近於真實分布，這是已知目前文獻中沒有使用過的估計分布方式。

再者使用離散事件模擬驗證本研究之結果，首先驗證長短期記憶神經網路識別未知環境之準確度，將估計資料視覺化後便可以知道其估計能力相當良好。接下來透過動態規劃求解訂價得到在已知環境下之真實利潤，並與模擬驗證得到之利潤計算偏好學習成本 Gap，得知使用本研究之方法學習真實利潤所需要花費的成本非常少，在面對不同分布情況下都能限制於 5% 內。

6.2 未來研究方向

本研究目前只假設消費者支付意願分布為常態/伽瑪分布，然而在實際市場環境中應為服從任意分布，因此未來可以嘗試更多種分布下市場之銷售情形；此外，除了 KL Divergence 之外還有很多計算分布差異的方式，未來也可以使用不

同方式計算 Loss；最後，本研究是以航空公司販售單一航班之機票為例，屬於簡單型動態規劃問題，未來可以擴展成更為複雜之動態規劃問題。



參考文獻



朱婉琳(2019). 利用人工智慧之消費者偏好學習與動態訂價方法. 臺灣大學工業工程學研究所學位論文.

Yang, Yang, Wan-Ling Chu, and Cheng-Hung Wu. "Learning customer preferences and dynamic pricing for perishable products." *Computers & Industrial Engineering* 171 (2022): 108440.

den Boer, Arnoud V., and Bert Zwart. "Dynamic pricing and learning with finite inventories." *Operations research* 63.4 (2015): 965-978.

Lin, Kyle Y. "Dynamic pricing with real-time demand learning." *European Journal of Operational Research* 174.1 (2006): 522-538.

Gallego, Guillermo, and Garrett Van Ryzin. "Optimal dynamic pricing of inventories with stochastic demand over finite horizons." *Management science* 40.8 (1994): 999-1020.

KINCAID, WM, and DA DARLING. An inventory pricing problem. MICHIGAN UNIV ANN ARBOR, 1962.

Popescu, Ioana, and Yaozhong Wu. "Dynamic pricing strategies with reference effects." *Operations research* 55.3 (2007): 413-429.

Caro, Felipe, and Jérémie Gallien. "Dynamic assortment with demand learning for seasonal consumer goods." *Management science* 53.2 (2007): 276-292.

Bellman, Richard. "The theory of dynamic programming." *Bulletin of the American Mathematical Society* 60.6 (1954): 503-515.

Li, Duan, et al. "Mitigation of curse of dimensionality in dynamic programming." *IFAC Proceedings Volumes* 41.2 (2008): 7778-7783.



Ghose, Tapu Kumar, and Thomas T. Tran. "Dynamic pricing in electronic commerce using neural network." E-Technologies: Innovation in an Open World: 4th International Conference, MCETECH 2009, Ottawa, Canada, May 4-6, 2009. Proceedings 4. Springer Berlin Heidelberg, 2009.

Ghose, Tapu Kumar, and Thomas T. Tran. "A dynamic pricing approach in e-commerce based on multiple purchase attributes." Advances in Artificial Intelligence: 23rd Canadian Conference on Artificial Intelligence, Canadian AI 2010, Ottawa, Canada, May 31–June 2, 2010. Proceedings 23. Springer Berlin Heidelberg, 2010.

AmalNick, M., and Roozbeh Qorbanian. "Dynamic pricing using wavelet neural network under uncertain demands." Decision Science Letters 6.3 (2017): 251-260.

Kong, Danxia. "One dynamic pricing strategy in agent economy using neural network based on online learning." IEEE/WIC/ACM International Conference on Web Intelligence (WI'04). IEEE, 2004.

Genser, Alexander, and Anastasios Kouvelas. "Dynamic optimal congestion pricing in multi-region urban networks by application of a Multi-Layer-Neural network." Transportation Research Part C: Emerging Technologies 134 (2022): 103485.

Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. "Learning representations by back-propagating errors." nature 323.6088 (1986): 533-536.

Schäfer, Anton Maximilian, and Hans Georg Zimmermann. "Recurrent neural networks are universal approximators." Artificial Neural Networks–ICANN 2006: 16th International Conference, Athens, Greece, September 10-14, 2006. Proceedings, Part I 16. Springer Berlin Heidelberg, 2006.

Bengio, Yoshua, Patrice Simard, and Paolo Frasconi. "Learning long-term dependencies with gradient descent is difficult." IEEE transactions on neural networks 5.2 (1994): 157-166.

Pascanu, Razvan, Tomas Mikolov, and Yoshua Bengio. "On the difficulty of training recurrent neural networks." International conference on machine learning. Pmlr, 2013.

Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9.8 (1997): 1735-1780.

Chung, Junyoung, et al. "Empirical evaluation of gated recurrent neural networks on sequence modeling." arXiv preprint arXiv:1412.3555 (2014).

Nelson, David MQ, Adriano CM Pereira, and Renato A. De Oliveira. "Stock market's price movement prediction with LSTM neural networks." 2017 International joint conference on neural networks (IJCNN). Ieee, 2017.

Sun, Yonghui, et al. "Ultra short-term probability prediction of wind power based on LSTM network and condition normal distribution." Wind Energy 23.1 (2020): 63-76.

Li, Qingliang, et al. "GANs-LSTM model for soil temperature estimation from meteorological: a new approach." IEEE Access 8 (2020): 59427-59443.

Yu, Shuyang, et al. "A domain adaptive convolutional LSTM model for prognostic remaining useful life estimation under variant conditions." 2019 Prognostics and System Health Management Conference (PHM-Paris). IEEE, 2019.

Sundermeyer, Martin, Hermann Ney, and Ralf Schlüter. "From feedforward to recurrent LSTM neural networks for language modeling." IEEE/ACM Transactions on Audio, Speech, and Language Processing 23.3 (2015): 517-529.

DeShon, Richard P. "A generalizability theory perspective on measurement error corrections in validity generalization." Validity generalization: A critical review (2003): 365-402.

Lipton, Zachary C., et al. "Learning to diagnose with LSTM recurrent neural networks." arXiv preprint arXiv:1511.03677 (2015).

Ji, Hongquan, Xiao He, and Donghua Zhou. "Diagnosis of sensor precision degradation using Kullback-Leibler divergence." *The Canadian Journal of Chemical Engineering* 96.2 (2018): 434-443.

Delpha, Claude, Demba Diallo, and Abdulrahman Youssef. "Kullback-Leibler Divergence for fault estimation and isolation: Application to Gamma distributed data." *Mechanical Systems and Signal Processing* 93 (2017): 118-135.

Kullback, Solomon, and Richard A. Leibler. "On information and sufficiency." *The annals of mathematical statistics* 22.1 (1951): 79-86.

Rust, John. "Has dynamic programming improved decision making?." *Annual Review of Economics* 11 (2019): 833-858.

Escobari, Diego. "Dynamic pricing, advance sales and aggregate demand learning in airlines." *The Journal of Industrial Economics* 60.4 (2012): 697-724.

Zhao, Wen, and Yu-Sheng Zheng. "Optimal dynamic pricing for perishable assets with nonhomogeneous demand." *Management science* 46.3 (2000): 375-388.

Shewalkar, Apeksha, Deepika Nyavanandi, and Simone A. Ludwig. "Performance evaluation of deep neural networks applied to speech recognition: RNN, LSTM and GRU." *Journal of Artificial Intelligence and Soft Computing Research* 9.4 (2019): 235-245.

Cahuantzi, Roberto, Xinye Chen, and Stefan Güttel. "A comparison of LSTM and GRU networks for learning symbolic sequences." *Science and Information Conference*.

Cham: Springer Nature Switzerland, 2023.

Astawa, I. Nyoman Gede Arya, I. Putu Bagus Arya Pradnyana, and I. Ketut Suwintana. "Comparison of RNN, LSTM, and GRU Methods on Forecasting Website Visitors." *Journal of Computer Science and Technology Studies* 4.2 (2022): 11-18.

Mahjoub, Sameh, et al. "Predicting Energy Consumption Using LSTM, Multi-Layer GRU and Drop-GRU Neural Networks." *Sensors* 22.11 (2022): 4062.

Liu, Zhaoyang, et al. "Context-aware attention LSTM network for flood prediction." 2018 24th international conference on pattern recognition (ICPR). IEEE, 2018.

Kutschinski, Erich, Thomas Uthmann, and Daniel Polani. "Learning competitive pricing strategies by multi-agent reinforcement learning." *Journal of Economic Dynamics and Control* 27.11-12 (2003): 2207-2218.

Mullen, Patrick B., et al. "Particle swarm optimization in dynamic pricing." 2006 IEEE International Conference on Evolutionary Computation. IEEE, 2006.

Lu, Renzhi, Seung Ho Hong, and Xiongfeng Zhang. "A dynamic pricing demand response algorithm for smart grid: Reinforcement learning approach." *Applied energy* 220 (2018): 220-230.

Maestre, Roberto, et al. "Reinforcement learning for fair dynamic pricing." *Intelligent Systems and Applications: Proceedings of the 2018 Intelligent Systems Conference (IntelliSys) Volume 1*. Springer International Publishing, 2019.

Schwind, Michael, and Oliver Wendt. "Dynamic pricing of information products based on reinforcement learning: A yield-management approach." *Annual Conference on Artificial Intelligence*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002.

Su, Xuanming. "Intertemporal pricing with strategic customer behavior." *Management Science* 53.5 (2007): 726-741.

Aviv, Yossi, and Amit Pazgal. "Optimal pricing of seasonal products in the presence of forward-looking consumers." *Manufacturing & service operations management* 10.3 (2008): 339-359.

Chade, Hector, and Virginia Vera De Serio. "Pricing, learning, and strategic behavior in a single-sale model." *Economic Theory* 19 (2002): 333-353.



Mersereau, Adam J., and Dan Zhang. "Markdown pricing with unknown fraction of strategic customers." *Manufacturing & Service Operations Management* 14.3 (2012): 355-370.

Kachani, Soulaymane, Georgia Perakis, and Carine Simon. "Modeling the transient nature of dynamic pricing with demand learning in a competitive environment." *Network science, nonlinear science and infrastructure systems* (2007): 223-267.

Priester, Anna, Thomas Robbert, and Stefan Roth. "A special price just for you: Effects of personalized dynamic pricing on consumer fairness perceptions." *Journal of Revenue and Pricing Management* 19 (2020): 99-112.

Shakya, Siddhartha, Fernando Oliveira, and Gilbert Owusu. "Analysing the effect of demand uncertainty in dynamic pricing with EAs." *International Conference on Innovative Techniques and Applications of Artificial Intelligence*. London: Springer London, 2008.

Shakya, Siddhartha, et al. "Neural network demand models and evolutionary optimisers for dynamic pricing." *Knowledge-Based Systems* 29 (2012): 44-53.

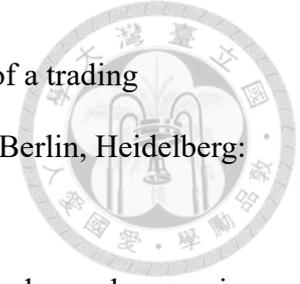
Teodorović, Dušan, and Praveen Edara. "A real-time road pricing system: The case of a two-link parallel network." *Computers & Operations Research* 34.1 (2007): 2-27.

Cournot, Antoine Augustin. *Researches into the Mathematical Principles of the Theory of Wealth*. New York: Macmillan Company, 1927 [c1897], 1927.

Mills, Edwin S. "Uncertainty and price theory." *The Quarterly Journal of Economics* 73.1 (1959): 116-130.

Nevins, Arthur J. "Some effects of uncertainty: Simulation of a model of price." *The Quarterly Journal of Economics* 80.1 (1966): 73-87.

Sandmo, Agnar. "On the theory of the competitive firm under price uncertainty." *The American Economic Review* 61.1 (1971): 65-73.



Aoki, Masanao. "On a dual control approach to the pricing policies of a trading specialist." IFIP Technical Conference on Optimization Techniques. Berlin, Heidelberg: Springer Berlin Heidelberg, 1973.

Qu, Huashuai, Ilya O. Ryzhov, and Michael C. Fu. "Learning logistic demand curves in business-to-business pricing." 2013 Winter Simulations Conference (WSC). IEEE, 2013.

Wruck, Eric Gordon. Dynamic pricing implications of uncertainty about demand. Cornell University, 1989.

Venezia, Itzhak. "Optimal investments in market research." *European Journal of Operational Research* 18.2 (1984): 198-207.

Morales-Enciso, Sergio, and Jürgen Branke. "Revenue maximization through dynamic pricing under unknown market behaviour." 3rd Student Conference on Operational Research. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2012.

Carvalho, Alexandre Xavier Ywata de, and Martin L. Puterman. "Dynamic optimization and learning: How should a manager set prices when the demand function is unknown?." (2005).

Eren, Serkan S., and Costis Maglaras. "Monopoly pricing with limited demand information." *Journal of revenue and pricing management* 9 (2010): 23-48.

Farias, Vivek F., and Benjamin Van Roy. "Dynamic pricing with a prior on market response." *Operations Research* 58.1 (2010): 16-29.

Avramidis, Athanassios. "Learning in revenue management: exploiting estimation of arrival rate and price response." Pre-print (2011).

Chen, Yongmin, and Ruqu Wang. "Learning buyers' valuation distribution in posted-price selling." *Economic Theory* 14 (1999): 417-428.

Aviv, Yossi, and Amit Pazgal. "A partially observed Markov decision process for dynamic pricing." *Management science* 51.9 (2005): 1400-1416.

Xiong, Yu, Gendao Li, and Kiran Jude Fernandes. "Dynamic pricing model and algorithm for perishable products with fuzzy demand." *Applied Stochastic Models in Business and Industry* 26.6 (2010): 758-774.

Abdella, Juhar Ahmed, et al. "Airline ticket price and demand prediction: A survey." *Journal of King Saud University-Computer and Information Sciences* 33.4 (2021): 375-391.

Hole, Arne Risa, and Julie Riise Kolstad. "Mixed logit estimation of willingness to pay distributions: a comparison of models in preference and WTP space using data from a health-related choice experiment." *Empirical Economics* 42 (2012): 445-469.

Chen, Shiyu, et al. "Dynamic pricing for smart mobile edge computing: A reinforcement learning approach." *IEEE Wireless Communications Letters* 10.4 (2020): 700-704.

Wang, Yining, et al. "Uncertainty quantification for demand prediction in contextual dynamic pricing." *Production and Operations Management* 30.6 (2021): 1703-1717.

Keskin, N. Bora, Yuexing Li, and Jing-Sheng Song. "Data-driven dynamic pricing and ordering with perishable inventory in a changing environment." *Management Science* 68.3 (2022): 1938-1958.

Asghari, Mohammad, and Cyrus Shahabi. "An on-line truthful and individually rational pricing mechanism for ride-sharing." *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. 2017.

Asghari, Mohammad. *Dynamic Pricing and Task Assignment in Real-Time Spatial Crowdsourcing Platforms*. Diss. University of Southern California, 2018.



Javanmard, Adel, Hamid Nazerzadeh, and Simeng Shao. "Multi-product dynamic pricing in high-dimensions with heterogeneous price sensitivity." 2020 IEEE International Symposium on Information Theory (ISIT). IEEE, 2020.

Li, Yanbin, et al. "Dynamic pricing based electric vehicle charging station location strategy using reinforcement learning." *Energy* 281 (2023): 128284.

Xiao, Chunjing, et al. "Modeling Behavior and Attribute feedback based Flight Recommendation for Dynamic Pricing." 2022 IEEE Smartworld, Ubiquitous Intelligence & Computing, Scalable Computing & Communications, Digital Twin, Privacy Computing, Metaverse, Autonomous & Trusted Vehicles (SmartWorld/UIC/ScalCom/DigitalTwin/PriComp/Meta). IEEE, 2022.

Wen, Chieh-Hua, and Po-Hung Chen. "Passenger booking timing for low-cost airlines: A continuous logit approach." *Journal of Air Transport Management* 64 (2017): 91-99.

Escobari, Diego. "Estimating dynamic demand for airlines." *Economics Letters* 124.1 (2014): 26-29.

An, Bo, et al. "Data-driven frequency-based airline profit maximization." *ACM Transactions on Intelligent Systems and Technology (TIST)* 8.4 (2017): 1-28.

Yuan, Hui, Wei Xu, and Chengfu Yang. "A user behavior-based ticket sales prediction using data mining tools: An empirical study in an OTA company." 2014 11th International Conference on Service Systems and Service Management (ICSSSM). IEEE, 2014.

Mostafaeipour, Ali, Alireza Goli, and Mojtaba Qolipour. "Prediction of air travel demand using a hybrid artificial neural network (ANN) with Bat and Firefly algorithms: a case study." *The journal of supercomputing* 74 (2018): 5461-5484.

Pan, Boxiao, et al. "A novel LSTM-based daily airline demand forecasting method using vertical and horizontal time series." *Trends and Applications in Knowledge*

Discovery and Data Mining: PAKDD 2018 Workshops, BDASC, BDM, ML4Cyber, PAISI, DaMEMO, Melbourne, VIC, Australia, June 3, 2018, Revised Selected Papers

22. Springer International Publishing, 2018.

Kumar, Avinash, Sobhangi Sarkar, and Chittaranjan Pradhan. "Malaria disease detection using cnn technique with sgd, rmsprop and adam optimizers." *Deep learning techniques for biomedical and health informatics* (2020): 211-230.

Srivastava, Nitish, et al. "Dropout: a simple way to prevent neural networks from overfitting." *The journal of machine learning research* 15.1 (2014): 1929-1958.

Nwankpa, Chigozie, et al. "Activation functions: Comparison of trends in practice and research for deep learning." *arXiv preprint arXiv:1811.03378* (2018).

Chen, Yuwen, et al. "An ensemble learning based approach for building airfare forecast service." *2015 IEEE International Conference on Big Data (Big Data)*. IEEE, 2015.

