Purposeful underestimation of demands for the airline seat allocation with incomplete information

Lijian Chen*

School of Business Administration,
Department of MIS, Operations Management, and Decision Sciences,
University of Dayton,
Dayton, OH 45469, USA
Email: lchen1@udayton.edu
*Corresponding author

Dengfeng Sun

School of Aeronautics and Astronautics Engineering,
Purdue University,
West Lafayette, IN 47907-2045, USA
Email: dsun@purdue.edu

Wen-Chyuan Chiang

Collins College of Business,
Tulsa, Oklahoma 74104-9700, USA
Email: wen-chiang@utulsa.edu

Shuguang He

School of Management,
Tianjin University,
Tianjin 300072, China
Email: shuguanghe@tju.edu.cn

Abstract: We study stochastic programming formulations for the origin-destination model in airline seat allocation under uncertainty. In particular, we focus on solving the stability issues of the traditional probabilistic model by purposefully underestimating the demands. The stochastic seat allocation models assume at least the possession of the distributional information, which is usually difficult to satisfy in a constantly changing environment. We propose a heuristic that consists of dynamically incorporating available information by solving a sequence of stochastic programming models. We show that the proposed method, named 'seat reservation (SR)', can ease most negative effects of incomplete distributional information and under some restrictive conditions, the SR will yield optimal revenue. The seat reservation method suggests that a revenue management company must (1) obtain timely results using adequately up-to-date computational facilities; (2) be conservative when allocating resources and (3) actively and continually revise previous estimations.

Keywords: stochastic programming; incomplete distributional information; airline seat allocation; revenue management.

Reference to this paper should be made as follows: Chen, L., Sun, D., Chiang, W-C. and He, S. (2014) 'Purposeful underestimation of demands for the airline seat allocation with incomplete information', *Int. J. Revenue Management*, Vol. 8, No. 1, pp.34–55.

Biographical notes: Lijian Chen is now an Assistant Professor at School of Business Administration, University of Dayton. He received his PhD from the Ohio State University in 2006. Since graduation, he has been working in revenue management industry and academia. He publishes 10+ business journal articles on journals of *Production and Operations Management*, *European Journal of Operational Management*, and *Annals of Operations Research*.

Dengfeng Sun is now an Assistant Professor at School of Aeronautics and Astronautics Engineering, Purdue University. He received his PhD from the University of California, Berkeley in 2009. He has been working in NASA before becoming a faculty member in Purdue University. His research is funded by the National Science Foundation.

Wen-Chyuan Chiang is Professor of Operations Management at Collins College of Business Administration, The University of Tulsa. He received his PhD in Operations Management and his MBA from the University of Texas at Austin and his BA from the National Taiwan Normal University. His research interests include revenue management, supply chain management, distribution and logistics, manufacturing operations, and AI applications to operations management problems. His research appears in *Management Science*, *International Journal of Revenue Management*, *IIE Transactions*, *Decision Sciences*, *International Journal of Production Research, European Journal of Operational Research*, *Annals of Operations Research*, *Journal of the Operational Research Society* and *INFORMS Journal on Computing*, among others. He has been the recipient of the Mayo Excellence in Research Award six times. He is Editor of *International Journal of Revenue Management* and a Senior Editor of Production and Operations Management.

Shuguang He is now an Associate Professor at the School of Management, Tianjin University, PRChina. He received his PhD on Operations Management in 2001. He visited Virginia Tech in 2001 and since then, he has published numerous business papers on transportation pricing, revenue management, management information system, and operations management.

1 Introduction

Revenue management involves the application of quantitative techniques to improve profits by controlling the prices and availabilities of various products that are produced with scarce resources. The best-known and earliest revenue management application occurred in the airline industry, where the products are tickets and the resources are seats on flights. With many successful applications, revenue management has received considerable attention in the past few years from both practitioners and academics. For example, many industries (e.g., car rental, hotel, travel agencies, and automakers) are actively using revenue management methodologies to increase their operational revenues. Among these methodologies on revenue management, there are two major categories: quantity-based and price-based revenue management systems. The representative industry for quantity-based revenue management is the airline industry. Other industries, such as retail, and hotel, practice price-based revenue management. In this paper, we focus on the revenue management practice at the airline industry and our model is thus for the quantity-based seat allocation.

Airline seat allocation is commonly depicted as a sequential decision problem over a fixed time horizon in which one decides whether each request for a ticket should be accepted or rejected. A typical assumption is that one can separate the demand for individual itinerary-class pairs, i.e., that each request is from a particular class on a particular itinerary and yields a predetermined fare level. Typically, a class is determined by particular constraints that are associated with the ticket rather than the physical seat. The classic example involves customers traveling for leisure and those traveling on business. The former group typically books in advance and is more price-sensitive; the latter group behaves in the opposite way. Airline companies attempt to sell as many seats as possible to high-fare paying customers to counter the potential loss that results from unsold seats. In most cases, rejecting an early request saves the seat for a later booking, which is likely to be high-fare request; at the same time, however, that action creates the risk of flying with empty seats. On the other hand, although accepting early requests raises the percentage of occupancy, it creates the risk of rejecting a future high-fare request due to constraints on capacity. Many airline companies adopt overbooking to cancel the impact of demand uncertainty (e.g., Aydin et al. (2012)).

There are many models to address the airline seat allocation problem and neither of them is perfect. When the customer arrival is modelled as many sequential events, the model would be dynamic and the solution of the model would be a control policy to maximise revenue. When the customer demands are modelled as aggregated lump sum by booking classes, a mathematical programming model (static allocation model or static model) will be adopted and the sequential factor of the customer arrival is ignored. The dynamic model seems to be superior in terms of performance because the static model solution is known to be sub-optimal that the method is restricted to partitioned booking limits. For example, high-fare customers' booking requests may be denied when there are still seats available for low-fare booking classes. Nevertheless, the dynamic model is less likely to be modelled by a computationally tractable operations research model. Consider a 10-period booking process for two booking classes and there is at most 1 customer request during each period per class. The overall number of outcomes of the sequence of arrivals is $(2^2)^{10} = 148,576$, which is a large number for such a small problem.

In addition to modelling techniques, the airline seat allocation problem is greatly complicated by factors such as volatile, stochastic demand for air travel, customer behaviours, and diversion of passengers to buy-up or buy-down other classes. The static model becomes promising when the customer arrival is modelled as the total number of requests per class. The sub-optimality mentioned above can be, at least partially, remedied by various nesting heuristics as a post-optimisation treatment to implement the optimal solution of the static model for the airline seat allocation. Thus, the airline companies adopt the static model in practice because, first, the static model can be easily solved and implemented for a network of flights; and second, the dual information can be extracted at no additional cost and it would be used to derive the bid prices of all booking classes.

The static model has been adopted by many airlines such as American Airline and United Airline. In particular, when the demands are modelled as random vectors, the static model becomes a stochastic programming model (or stochastic seat allocation model) which is to maximise the revenue by average. This model is not perfect as well regardless the implementation of the nesting heuristic. The stochastic programming models are fundamentally based on the assumption that the random variables' distributional information is known. However, that assumption may be overly optimistic; in reality, the distributional information feed could be either incomplete or erroneous. When the real distribution differs from the underlying distribution that was previously estimated even if we considered a minor perturbation for a mathematical programming model as low as 0.1% the consequence could be a considerably different optimal solution (see Ben-Tal and Nemirovski (1999)). The stability issue of stochastic programming with incomplete distributional information is an important topic and very few methods have been developed to solve the stability issue in practice. The state-of-the-art of research focuses mainly on the Lipschitz property rather than developing practical methods to improve the model performance.

In this paper, we propose a heuristic to solve the stability issues of the stochastic seat allocation model. In our opinion, we need to acknowledge the fact that the existence of incomplete distributional information is common and mostly inevitable. In order to incorporate available information to the allocation planning, a sequence of estimations as well as a sequence of stochastic programs will be generated and solved. The booking policy will be updated accordingly. This idea is not new and many results have been presented to the academia. For example, Cooper (2002) and Chen and Homem-de-Mello (2010b) present the seat allocation model with periodically updating parameters. These results suggest that updating estimations will nearly always benefit the airline with higher expected revenue. In Lemke et al. (2013), the authors suggest that the demand, sometimes, can be forecasted in an evolving way.

The managerial implication of this heuristic is straightforward. Given the fact that the underlying distribution is different from the real distribution, the optimal solution with the underlying distribution may not be optimal in reality. We need to revise our seat allocation with information becoming available. If we can assume that our estimation of the underlying distribution eventually converges *weakly* to the real distribution, the allocation plan will converge to the true optimal. Of course, this assumption appears to be overly optimistic because the number of available seats is another contributing factor to the optimal solution. As the booking moves forward, the number of available seats may drop to a level by which the stochastic allocation model may never be optimal even if the demand estimation is correct.

Our heuristic mainly focuses on the effective control on the number of available seats because the updated estimations regarding the distributional parameters could be continuously re-estimated as more information becomes available. The later estimations are imposed with less randomness and thereby are expected to be more accurate. The number of available seats, however, are exposed to more uncertainties. For example, the entire coach class tickets may be sold out over night if there is an unexpected demand hike of low-fare classes and it may lead to the potential risk of losing high-fare class sales. Thus, instead of allocating all the resources once, our heuristic allocates seats to booking classes gradually and continually revise previous estimation. Because of this main focus, we name this method, *seat reservation* (SR) for the seat allocations with incomplete distributional information.

The airline seat allocation can be assumed to be risk-neutral because the airline needs to decide thousands of times per day whether or not to accept or reject customer booking requests. The objective is to manage the opening and closing of discount fare classes in such a way that the *expected* revenues are maximised (see Talluri and Van Ryzin (2005)). Thus, in this paper, we focus on the stochastic programming with a risk-neutral objective. There are exceptions. For example in Huang and Chang (2011) and Graf and Kimms (2011), the authors propose an alternative model by integrating risk to the seat allocation model. In Zhuang and Li (2011), the authors present a risk-averse model for a specific utility function. In practice, the non-airline companies may be more interested in a risk-averse revenue management model. In Ferrer et al. (2012), the authors suggest that the risk-averse model would play a central role in the retailing industry. When the objective is mean-variance (MV) based risk-averse, our SR heuristic will still function as expected.

The core contribution of this paper is a computationally tractable heuristic to tackle the stability issue of the stochastic programming for the airline revenue management. The idea of the SR heuristic is consistent with the common sense, "save for a rainy day". When the future is uncertain, we'd better to reserve resources for unexpected events. For the airline seat allocation, it is wise to reserve a certain amount of seats for ever-existing estimation errors rather than allocating all the seats at the beginning. At the same time, updating previous estimates would be actively pursued. The rest of the paper is organised as follows: In Section 2, we introduce the notation and describe stochastic models for the seat allocation problem. The SR heuristic is presented in detail and we also discuss the risk-averse model for the companies outside the airline industry in Section 3. We present the numerical results in Section 4. Concluding remarks are presented in Section 5.

2 Seat allocation methods

Following standard models in the literature, we consider a network of flights involving p booking classes of customers. The model can represent demand for a network of flights that depart on a particular date. Each customer requests one out of n possible itineraries, so we have $r := np$ itinerary-fare class combinations. The booking process is realised over a time horizon of length τ indexed by t. When t is discrete, $t \in \{1, 2, \dots, \tau\}$ and when t is continuous, we have $t \in [0, \tau]$. Let $\{N_{jk}(t)\}$ denote the booking process generated by the arrivals of class-k customers who request itinerary j. Typical cases in the revenue management literature are

- $\{N_{jk}(t)\}$ is a (possibly non-homogeneous) Poisson process
- there is at most one unit of demand per time period. We assume that arrival processes corresponding to different pairs (j, k) are independent of each other.

The demand for itinerary-class (j, k) over the entire horizon is denoted by ξ_{jk} (i.e., $\xi_{jk} = N_{jk}(\tau)$) and we denote by ξ the whole vector (ξ_{jk}). If ξ is a discrete random vector, let S_{jk} represent the number of possible values taken by ξ_{jk}. The network is comprised of m leg cabin combinations, with capacities $c := (c_1, \dots, c_m)$, and is represented by an $(m \times np)$-matrix $A := (a_{i,jk})$. The entry $a_{i,jk} \in \{0, 1\}$ indicates whether class k customers use leg i in itinerary j. Most policies we study within this paper are of allocation type. We denote by

x_{jk} the decision variable corresponding to the number of seats to be allocated to class k in itinerary j. Whenever an itinerary-class pair (j,k) is accepted, the revenue corresponding to the fare f_{jk} accrues. A customer's request is rejected if no seats are available for his or her itinerary-class; in such a case, no revenue is realised. The vectors of decision variables and fares are denoted respectively by $x = (x_{jk})$ and $f = (f_{jk})$.

Allocation methods require solving an optimisation problem to find the initial allocations before the booking process begins. The solution x^* becomes the allocation plan and the implementation of x^* is very simple, accepting at most x^*_{jk} class k customers in itinerary j. Notice that the policy is well-defined even if the solution x^* is not an integer.

2.1 Stochastic programming model

When the random demand ξ is in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω is the sample space and $\mathbb{P}$ is the underlying distribution of ξ with finite support, we need to make the following assumptions to ensure that the stochastic models is well-defined.

A The feasible region, $X := \{x : Ax \leq c, x \in \mathbb{R}^r_+\}$, is not empty

B The available resource vector c is deterministic and completely known.

The resulting stochastic allocation model is formulated into the stochastic programming with recourse as follows:

$$\begin{aligned} \nu^* := \max\ & f'x + \mathbb{E}[Q(x,\xi)] \\ \text{subject to:}\quad & Ax \leq c \\ & x \in \mathbb{R}^r_+ \end{aligned} \tag{SLP}$$

where $Q(x,\xi) := \max\{-f'y | x - y \leq \xi,\ y \geq 0\}$. Model (SLP) can be solved by several numerical techniques efficiently, regardless of the scale of the problem. With assumptions 2.1 and 2.1, the function $Q(x,\xi)$ is a concave function with respect to x and so is the objective function $f'x + \mathbb{E}[Q(x,\xi)]$ of model (SLP). For example, in Chen and Homem-de-Mello (2010b), the authors present that there are $O(Snp)$ variables and $O(Snp + m)$ constraints where $S := \max_{j,k} S_{jk}$ is the equivalent linear programming instance of model (SLP).

2.2 Review of the stability results of stochastic programming

In the real-world airline seat allocation problem, the underlying distribution of model (SLP) may be different from the real distribution due to the error of the estimation. In other words, model (SLP) is solved with underlying probability measure $\mathbb{P}$ while the real distribution is $\mathbb{Q}$. Consequently, the optimal solution obtained may be different from the true optimal solution. This problem is named the stability issue of the stochastic programming. In the literature, this problem is called the 'quantitative stability' which is usually studied by the probabilistic metric method. The quantitative stability refers primarily to continuity properties of the optimal objective value and a suitable distance (e.g., ζ-structure distance in Romisch and Schultz (1993) and Rachev and Römisch (2002)) between the underlying and real probability measures, e.g., ζ-structure distance, which is defined as follows:

$$d_{\mathcal{F}}(\mathbb{P},\mathbb{Q}) := \sup_{g\in\mathcal{F}} \left| \int_\Omega g(\xi)d\mathbb{P}(\xi) - \int_\Omega g(\xi)d\mathbb{Q}(\xi) \right| \tag{2.1}$$

where g is a measurable function in $\mathcal{F}$ and $\xi \in (\Omega, \mathcal{F}, \cdot)$. In the remaining part of the paper, we use "dist$(\mathbb{P}, \mathbb{Q})$" to represent the distance without differentiating the probabilistic metric being applied. We have the following result:

Theorem 1: *Let assumptions 2.1 and 2.1 be satisfied, then there exists a constant $L > 0$ such that*

$$|\nu^*(\mathbb{P}) - \nu^*(\mathbb{Q})| \leq L \times \mathit{dist}(\mathbb{P}, \mathbb{Q}) \tag{2.2}$$

where $\nu^*(\cdot)$ represents the optimal value of equation (SLP) *under corresponding distributions $\mathbb{P}$ and $\mathbb{Q}$, respectively.*

Proof: Please refer to (Ruszczynski and Shapiro, 2003, Chapter 8). □

Let $X^*(\mathbb{P})$ and $X^*(\mathbb{Q})$ be the *sets* of optimal solutions of equation (SLP) under the distributions $\mathbb{P}$ and $\mathbb{Q}$, respectively. By Rockafellar and Wets (1998), we have the following theorem:

Theorem 2: *Let assumptions 2.1 to 2.1 be satisfied and $dist(\mathbb{P}, \mathbb{Q}) < \infty$, there exists a constant $\hat{L}$ such that*

$$X^*(\mathbb{Q}) \subset \left(X^*(\mathbb{P}) \cup \hat{L} \times dist(\mathbb{P}, \mathbb{Q}) \times \mathbb{B} \right) \tag{2.3}$$

where $\mathbb{B}$ is an Euclidean unit ball.

Essentially, Theorems 1 and 2 show that the gap between two corresponding optimal values is bounded by two probability distributions' distance adjusted by a Lipschitz constant. Despite the promising theoretical results, the quantitative stability has quite limited impact in practice for two reasons. First, the real distribution $\mathbb{Q}$ is not available until it is revealed. Second, the Lipschitz constant is extremely difficult to evaluate, thereby constructing a tight bound becomes hardly possible.

Under some restrictive assumptions, a stochastic programming *without* perfect estimation on demands may still deliver a satisfactory performance that is *close* to the true optimal. To achieve such a goal, we need to dynamically incorporate available information to actively and continually revise the previous estimations and update the seat allocation plan. This idea is fundamentally backed by the following well known results in Nemirovski et al. (2009):

Definition 1: Let $\{t_s\}$ be the infinite sequence of time spots in $[0, \tau]$. The weak convergence of the sequence of probability measures $\{\mathbb{P}_{t_s}\}$ to $\mathbb{Q}$ means that for any bounded continuous function $g : \mathbb{R}^r \to \mathbb{R}$, we have

$$\int_{\mathbb{R}^r} g(\xi) d\mathbb{P}_{t_s}(\xi) \to \int_{\mathbb{R}^r} g(\xi) d\mathbb{Q}(\xi), \text{ as } s \to \infty. \tag{2.4}$$

The weak convergence will hold if and only if the sequence of probability distribution functions of $\mathbb{P}_{t_s}$ converges pointwise to the distribution function $F_{\mathbb{Q}}$ of $\mathbb{Q}$ at all continuity points of $F_{\mathbb{Q}}$.

Theorem 3: *Let $\{\mathbb{P}_{t_s}\}$ be a sequence of distributions that is weakly convergent to $\mathbb{Q}$. Then the sequence $\{\nu^*(\mathbb{P}_{t_s})\}$ converges to $\{\nu^*(\mathbb{Q})\}$; and*

$$\lim_{s\to\infty} \sup_{x\in X(\mathbb{P}_{t_s})} \mathit{dist}(x, X^*(\mathbb{Q})) = 0 \tag{2.5}$$

when $\mathcal{F}$ is uniformly integrable with respect to $\{\mathbb{P}_{t_s}\}$.

Our SR heuristic is derived from the above theorem. During a booking horizon $[0,\tau]$, we revisit our previous estimations at s time points $t_1,\ldots,t_s$ such that $t_1 \leq \ldots \leq t_s$. Let ℓ be the index of s such that $\ell = 1,2,\ldots,s$. At the time point t_ℓ, information prior to t_ℓ has been observed and we can forwardly incorporate it to pursue better estimations. Each 'estimation', $\mathbb{P}_{t_\ell}$ at time t_ℓ, is for the entire booking horizon even if the randomness prior to time t_ℓ has been revealed. Consider two successive estimations $\mathbb{P}_{t_\ell}$ and $\mathbb{P}_{t_{\ell+1}}$, we always have,

$$\text{dist}(\mathbb{P}_{t_\ell},\mathbb{Q}) \geq \text{dist}(\mathbb{P}_{t_{\ell+1}},\mathbb{Q}), \ell = 1,\ldots,s \tag{2.6}$$

If we assume that the sequence of gradually improving estimations $\{\mathbb{P}_{t_\ell}\}$ is weakly convergent to the real distribution $\mathbb{Q}$, by applying Theorem 3, the sequence of optimal values of model (SLP) will converge to $\nu^*(\mathbb{Q})$. The sequence of optimal solutions of models (SLP) will become closer to the true optimal solution as well. If the seat allocation decisions were *not irrevocable*, then the optimal values obtained from continually improving estimations converge to the performances of revenue management models. In practice, when a 'better' and later estimation is obtained, the company tends to update the seat allocation plan to pursue better performance rather than sticking to the obsolete plans.

However, there is a problem of updating the seat allocation plan with newly acquired information. For some booking classes, their allocations have been heavily sold so that the updated allocation plan cannot be exercised. For example, a booking class has 5 seats allocated according to the past estimation. The demand of this booking class happened to be stronger that all 5 seats were sold prior to next update. With a better estimation, the previous seat allocation plan is revised to be 3 seats. Obviously, the updated seat allocation plan will not be carried out. Thus, in order to pursue the best booking performance, the airline needs to avoid over-allocating seats which is usually caused by either underestimation or overestimation of demands.

2.3 Overestimation and underestimation

Let $\xi_{\mathbb{P}}$ be the underlying distribution from the probability space $(\Omega,\mathcal{F},\mathbb{P})$ while the real distribution is $\xi_{\mathbb{Q}}$ on the probability space $(\Omega,\mathcal{F},\mathbb{Q})$. $\mathbb{P}$ is the underlying distribution from estimation, and $\mathbb{Q}$ is the real distribution.

Definition 2: The overestimation of demand is $\mathbb{P}(\xi_{\mathbb{Q}} \leq \xi_{\mathbb{P}}) = 1$, $\text{Var}(\xi_{\mathbb{Q}}) = \text{Var}(\xi_{\mathbb{P}})$ and underestimation $\mathbb{P}(\xi_{\mathbb{Q}} \geq \xi_{\mathbb{P}}) = 1$, $\text{Var}(\xi_{\mathbb{Q}}) = \text{Var}(\xi_{\mathbb{P}})$.

Let $\nu^*(\mathbb{P})$ and $\nu^*(\mathbb{Q})$ denote the optimal values; $x^*(\mathbb{P})$ and $x^*(\mathbb{Q})$ represent any optimal solutions of equation (SLP) under distributions $\mathbb{P}$ and $\mathbb{Q}$, respectively. We pose the following proposition:

Proposition 4: *The overestimation (underestimation) will lead to* $\nu^*(\mathbb{P}) \geq \nu^*(\mathbb{Q})$ *(*$\nu^*(\mathbb{P}) \leq \nu^*(\mathbb{Q})$*).*

Proof: Suppose that the company overestimates $\xi_{\mathbb{Q}}$ by $\xi_{\mathbb{P}}$. Since $Q(x, \xi_{\mathbb{Q}})$ is a concave and non-decreasing function with respect to $\xi_{\mathbb{Q}}$, for every feasible x of equation (SLP), we have $Q(x, \xi_{\mathbb{P}}) \geq Q(x, \xi_{\mathbb{Q}})$ and thus $f'x + \mathbb{E}[Q(x, \xi_{\mathbb{P}})] \geq f'x + \mathbb{E}[Q(x, \xi_{\mathbb{Q}})]$. □

Proposition 4 illustrates an illusory projection on revenue when demands are either overestimated or underestimated. This result is consistent with the observation that when the airline overestimates (underestimates) the demands, the profit expectation will be accordingly greater (less). In this paper, we are concerned with the performance of revenue management system, i.e., $f'x^*(\mathbb{P}) + \mathbb{E}[Q(x^*(\mathbb{P}), \xi_{\mathbb{Q}})]$, $\xi_{\mathbb{Q}} \in (\Omega, \mathcal{F}, \mathbb{Q})$, the expected total revenue in reality. The optimal value $\nu^*(\mathbb{P})$ is not the real revenue but a wrongly projected revenue with underlying distribution $\mathbb{P}$. With the real distribution $\mathbb{Q}$, neither performance under overestimation or underestimation will outperform the optimal value of $\nu^*(\mathbb{Q})$ because, by the definition of $x^*(\mathbb{Q})$:

$$f'x^*(\mathbb{P}) + \mathbb{E}[Q(x^*(\mathbb{P}), \xi_{\mathbb{Q}})] \leq f'x^*(\mathbb{Q}) + \mathbb{E}[Q(x^*(\mathbb{Q}), \xi_{\mathbb{Q}})] \text{ when } x^*(\mathbb{P}) \neq x^*(\mathbb{Q}). \tag{2.7}$$

We identify a tendency that overestimation (underestimation) implies $x^*(\mathbb{P}) \geq x^*(\mathbb{Q})$ ($x^*(\mathbb{P}) \leq x^*(\mathbb{Q})$) for high-fare booking classes. To show this tendency, consider an extreme case that none of the capacity constraints $Ax \leq c$ is tight, $x^*(\mathbb{P}) \geq x^*(\mathbb{Q})$ ($x^*(\mathbb{P}) \leq x^*(\mathbb{Q})$) is true for all the booking classes. This is because that when the seat capacity constraints are not tight, equation (SLP) becomes a unconstrained optimisation problem within the neighbourhood of the optimal solution. Since the objective of equation (SLP) is concave and non-decreasing with respect to x, we have $x^*(\mathbb{P}) \geq x^*(\mathbb{Q})$. Even if some capacity constraint(s) is tight, such a tendency may be true for high-fare class allocations that is much more important than the low-fare class allocations in terms of the revenue generation (see Talluri and Van Ryzin (1998)). When $x^*(\mathbb{P}) \geq x^*(\mathbb{Q})$ ($x^*(\mathbb{P}) \leq x^*(\mathbb{Q})$), we call this decision for seat allocation among multiple booking classes the *over-allocation plan (under-allocation plan)*.

The heuristic in the next section, the SR method, will focus on how to minimise the chance of an over-allocation plan. Although over-allocation plans and under-allocation plans both contribute to the sub-optimality, an over-allocation plan may lead to over sell to the customers of some booking classes who tends to pay less and book earlier. We are less concerned with the under-allocation plan because it will be easily revisited by re-solving equation (SLP) with the best estimation available. The revisit of the seat allocation will be efficiently solved and the computational cost of each revisit per se is nearly negligible.

3 The seat reservation method

In this section, we propose a heuristic named 'seat reservation (SR)' to address the stability issue of the mathematical programming models for airline seat allocation under uncertainty. By utilising the SR method, we are able to determine the time to start rejecting booking requests adaptively; to establish the connections; and more importantly, to justify the optimality from the theoretical aspect.

The goal of the SR method is to pursue close-to-optimal seat allocation model performance with incomplete distributional information. To pursue this goal, we must first investigate the causes of sub-optimality. According to the analysis in Section 2.3, if we relax the capacity constraints, the company tends to allocate more seats when overestimating, or, to allocate fewer seats when underestimating. When the company is over-allocating for low-fare booking classes and the demands from those booking classes are indeed stronger than expected, the company sells seats that should have been allocated for high-fare classes. Likewise, when the company is under-allocating resources to high-fare booking classes and the high-fare class demands happen to be stronger, the company must immediately revise the allocation plan to avoid any loss of high-fare customers. Therefore, over-allocating resources should be avoided at all costs because the company may sell seats earlier than it should and because less flexibility exists to avoid later regret. It is fair to say that over-allocating to low-fare classes always leads to poor booking performances under incomplete distributional information.

In comparison to the over-allocation plan, the under-allocation plan is actually a good choice. Consider a company that underestimates the demand entirely. There will be a tendency to allocate fewer seats to certain booking classes and that may lead to early rejections. Rather than simply rejecting the customer requests from certain booking classes, the company may update the seat allocation plan based on a 'better' estimation. As long as the updates on the distributional information and its corresponding seat allocation plan are calculated within a timely manner, the customers may be unaware of the existence of such updates. If the updated allocation plan adjusts the number of seats for certain booking classes, a rejection can either be avoided or confirmed. Thus, a company would completely avoid over-allocation to low-fare customers by *purposefully* underestimating demands by a fixed percentage $\theta \in (0, 1)$. Whenever a rejection is about to happen, the company will collect newly emerging information to adjust the allocation plan if necessary. Based on the idea of those purposeful underestimations, we propose the SR method as the solution to the stability issue of airline seat allocation.

There are three major components for the SR heuristic. First, the distributional information must be actively, continually, and quickly updated. In the literature of airline revenue management, the uncertain demands are overwhelmingly modelled as a non-homogeneous Poisson random vector. To update $\xi_{\mathbb{P}_{t_\ell}}$ at time t_ℓ, we only need to partially revise the arrival rates for all the booking classes. The update on demand distribution will be partial because the demand up to t_ℓ has been realised and becomes deterministic. The update will mainly focus on the uncertain demand from t_ℓ to τ and the adjustment on arrival rates per se, is rather straightforward and simple. Second, the SR heuristic must justify the closeness of obtained optimal solution to the true optimal, at least with restrictive assumptions. This component is mainly based on Definition 1 and Theorem 3. We need to assume the sequence of estimations, $\{\mathbb{P}_{t_\ell}\}, \ell = 1, \ldots, s$, which weakly converges to the real distribution $\mathbb{Q}$ as $s \to \infty$. That is, the sequence of probability distribution functions of $\mathbb{P}_{t_\ell}$ converges pointwise to the distribution function $F_{\mathbb{Q}}$ of $\mathbb{Q}$ at all continuity points of $F_{\mathbb{Q}}$. Third, it requires to solve the model (SLP) within a timely manner. This component is less of a concern because there are many algorithms and software packages to solve large scale stochastic programming.

3.1 Discussion on seat reservation

We now present the SR heuristic. We outline the SR heuristic into two phases: the phase to avoid rejections for all booking classes (Phase I in a short form) and the phase to avoid rejections for high-fare booking classes (Phase II in a short form).

3.1.1 Phase I of the seat reservation

By Belobaba and Weatherford (1996), low-fare customers tend to book earlier. Although it is true that the low-fare customers contribute less revenue comparing with the high-fare customers, the airline company will be reluctant to reject customers at the early booking as long as there is no over-allocation to the low-fare booking classes. Thus, it is critical to identify over-allocation to low-fare booking classes.

We avoid over-allocation to low-fare classes with the following procedure. Suppose the airline company update its demand estimation at time $t_1, \ldots, t_\ell, \ldots, t_s$ with distributions $\mathbb{P}_{t_1}, \ldots, \mathbb{P}_{t_s}$ and the real distribution is $\mathbb{Q}$. Let ξ_{t_ℓ} represent the *realised* demand till time t_ℓ and let $\Delta\xi_{t_\ell} := \xi_{t_\ell+1} - \xi_{t_\ell}$ which means the demand between time points t_ℓ and $t_{\ell+1}$. The number of booking requests that are received from 0 to t_ℓ is $y_{t_\ell} \in \mathbb{R}^r$. Consider the $(\ell-1)^{th}$ estimate of demand for the entire booking horizon would be $\mathbb{P}_{t_{\ell-1}}$ at time $t_{\ell-1}$ and the current optimal seat allocation would be $x^*(\mathbb{P}_{t_{\ell-1}})$. If the airline company does not allow any rejection to occur up to time t_ℓ, the following condition should be satisfied

$$y_{t_\ell} \leq x^*(\mathbb{P}_{t_{\ell-1}}) \qquad \text{(no rejection phase I)}$$

Definition 3: Let t_l be the earliest time point in $[0, \tau]$ such that $y_{t_l} \nleq x^*(\mathbb{P}_{t_l})\}$ which indicates that rejections are required.

There will be rejections and it is unrealistic to avoid any rejection for all the booking classes. The rejections for certain booking classes might happen when the amount of seats is sold out. Thus, the booking horizon would be divided into two phases. The phase I is $[0, t_l)$ in which no rejection to any booking class is necessary. The phase II is $[t_l, \tau]$ and only high-fare rejections are closely monitored. We must remark that the value of t_l is determined by the sample path rather than pre-determined.

In order to avoid over-allocation to low-fare classes, we need a predetermined parameter $\theta \in [0, 1]$ and we solve model (SLP) with demands ξ and $\theta\xi$ to obtain two optimal solutions $x^*(\mathbb{P}_{t_{\ell-1}})$ and $x^*_\theta(\mathbb{P}_{t_{\ell-1}}) :=$ argmin$\{f'x + \mathbb{E}[Q(x, \theta\xi)] | Ax \leq c, x \geq 0\}$. Consider any time t_ℓ in the phase I. If we have

$$y_{t_\ell} \leq \min\{x^*(\mathbb{P}_{t_{\ell-1}}), x^*_\theta(\mathbb{P}_{t_{\ell-1}})\}, \tag{3.1}$$

all booking requests will be honored. However, if equation (3.1) fails at time t_ℓ, we need to update the underlying distribution of ξ to $\mathbb{P}_{t_\ell}$. Then we test the following condition

$$y_{t_\ell} \leq \min\{x^*(\mathbb{P}_{t_\ell}), x^*_\theta(\mathbb{P}_{t_\ell})\}. \tag{3.2}$$

If equation (3.2) is satisfied, it indicates that the airline company adjusts the previous seat allocation to avoid any rejection by updating demand estimates. Otherwise, the phase I booking is over.

During the rejection-free phase I, we still frequently update seat allocations until the time t_l which indicates the end of phase I and rejections to low-fare classes would be necessary. The phase I SR can be summarised as follows.

Step 1-1. Choose $\theta > 0$ for the phase I.

Step 1-2. Let $\ell = 1$ and estimate the distribution of ξ as $\mathbb{P}_{t_\ell}$.

Step 1-3. Solve model (SLP) with demands ξ and $\theta\xi$ to obtain two optimal solutions $x^*(\mathbb{P}_{t_\ell})$ and $x^*_\theta(\mathbb{P}_{t_\ell}) :=$ argmin$\{f'x + \mathbb{E}[Q(x, \theta\xi)] | Ax \leq c, x \geq 0\}$.

Step 1-4. Let $\min\{x^*_\theta(\mathbb{P}_{t_\ell}), x^*(\mathbb{P}_{t_\ell})\} > \mathbf{0}$ be the seat allocation plan and continue booking until an incoming booking request is placed on a booking class with zero seats left. Then let $\ell := \ell + 1$ and update $\mathbb{P}_{t_\ell}$. If $t_\ell = \tau$, terminate booking and exit.

Step 1-5. Solve model (SLP) with updated distribution $\mathbb{P}_{t_\ell}$ to obtain two optimal solutions $x^*(\mathbb{P}_{t_\ell})$ and $x^*_\theta(\mathbb{P}_{t_\ell})$. If the updated seat allocation still suggests reject certain booking request(s), then we terminate the phase I SR procedure and move into the phase II seat reservation procedure. Otherwise, go to Step 1-4.

3.1.2 Phase II of the seat reservation

The termination of phase I indicates that some rejections need to be made and the airline company tends to do it at the low-fare booking classes mostly. We use $z_{t_\ell} \in \mathbb{R}^r$ to represent the number of rejections in the phase II up to time t_ℓ. Let $\mathcal{H}$ represent the set of the high-fare booking classes. We define a vector $H = [h_{11}; \ldots; h_{jk}; \ldots; h_{np}]' \in \mathbb{R}^{np}$ where $h_{jk} \in \{0, 1\}$ indicates whether class k customer in itinerary j is high-fare. We define the operator "$\cdot$" for two vectors $a, b \in \mathbb{R}^n$, $a = [a_1; \ldots; a_n]'$, $b = [b_1; \ldots; b_n]'$ as follows,

$$a \cdot b = [a_1 b_1; \ldots; a_n b_n]'$$

Thus, the number of high-fare booking requests up to time $t_\ell \in [t_l, \tau]$ is denoted as $y_{t_\ell} \cdot H$.

At phase II, the airline company focuses on the sales of high-fare customers. Since rejections of low-fare classes are inevitable in phase II, the goal of seat allocation is to reserve seats for the high-fare customers by mostly rejecting low-fare customers. Once a rejection to high-fare customer is about to be made, the airline company will update its allocation policy to either adjust the allocation to avoid such a rejection or confirm that such a rejection is necessary. In order to control the risk of over-allocation to low-fare classes, at phase II, the parameter $\theta \in [0, 1]$ inherited from the phase I will be used to purposefully underestimate the demands of high-fare classes. The number of booking requests is y_{t_ℓ} and the current allocation plan is $\min\{x^*(\mathbb{P}_{t_{\ell-1}}), x^*_\theta(\mathbb{P}_{t_{\ell-1}})\}$. This allocation plan is acquired by solving model (SLP) with demands ξ and $\theta\xi$ to obtain two optimal solutions $x^*(\mathbb{P}_{t_{\ell-1}})$ and $x^*_\theta(\mathbb{P}_{t_{\ell-1}}) :=$ argmin$\{f'x + \mathbb{E}[Q(x, \theta\xi)] | Ax \leq c, x \geq 0\}$. We will continue booking as long as the following condition is satisfied.

$$(y_{t_\ell} - z_{t_\ell}) \cdot H \leq \min\{x^*(\mathbb{P}_{t_{\ell-1}}), x^*_\theta(\mathbb{P}_{t_{\ell-1}})\} \cdot H \tag{3.3}$$

When equation (3.3) fails, it indicates that a rejection is about to be made. The airline company will update the estimate of ξ's distribution to $\mathbb{P}_{t_\ell}$ to see if

$$(y_{t_\ell} - z_{t_\ell}) \cdot H \leq \min\{x^*(\mathbb{P}_{t_\ell}), x^*_\theta(\mathbb{P}_{t_\ell})\} \cdot H \tag{3.4}$$

is satisfied. When equation (3.4) is satisfied, it indicates that the airline adjusts previous seat allocation by allocating more seats to high-fare classes. Otherwise, this rejection on a high-fare booking class is necessary due to the very low available inventory level. The rejections are more likely to happen to the low-fare booking classes at first and at the very end of booking, rejections of some high-fare booking classes are likely regardless how often we update the estimation of demands. The phase II seat reservation procedure is outlined as follows.

Step 2-1. Formulate equation (SLP) and define the set of the high-fare booking classes $\mathcal{H}$.

Step 2-2. Solve model (SLP) with demands ξ and $\theta\xi$ to obtain two optimal solutions $x^*(\mathbb{P}_{t_\ell})$ and $x^*_\theta(\mathbb{P}_{t_\ell}) :=$ argmin$\{f'x + \mathbb{E}[Q(x, \theta\xi)]|Ax \le c, x \ge 0\}$.

Step 2-3. Let $\min\{x^*_\theta(\mathbb{P}_i), x^*(\mathbb{P}_i)\} > \mathbf{0}$ be the seat allocation plan and continue booking until an incoming booking request is placed on a booking class with zero seats left. Then let $\ell := \ell + 1$ and update $\mathbb{P}_{t_\ell}$. If $t_\ell = \tau$, terminate booking and exit.

Step 2-4. Solve model (SLP) with updated distribution $\mathbb{P}_{t_\ell}$ to obtain two optimal solutions $x^*(\mathbb{P}_{t_\ell})$ and $x^*_\theta(\mathbb{P}_{t_\ell})$. If the updated seat allocation still suggests reject certain booking request(s), then we confirm the rejection of pending and upcoming booking requests associated with certain booking classes and go to Step 2-2. Otherwise, go to Step 2-3.

We discuss the practical perspective of the SR heuristic as a solution to the stability issue of seat allocation systems. Beyond the technical argument about the optimality, we find that the proposed SR method suggests three most important capabilities for the revenue management system with incomplete distributional information: conservative decisions, active and continual improvement on estimations, and an adequately up-to-date computational facility.

When the revenue management company encounters incomplete or erroneous information, any decision about the seat allocation should be made gradually. Otherwise, there will be very limited flexibility for the company to make high-fare sales. Each potential rejection for all booking classes in phase I and each potential rejection for high-fare classes will trigger the event of revising the previous estimation and updating allocation plans accordingly. Every potential rejection, in particular every high-fare booking request, will have a second chance, although the customer may not be aware of it. Being conservative about rejections and seat allocation will buy some time to obtain 'better' estimation and yield more mature decisions about rejections.

The company must have an adequately up-to-date computational facility to support the idea of conservative decision. According to the SR heuristic, the company must revise previous estimations and repeatedly solve large-scale stochastic programs with different distributions. For every potential rejection, the events of revising the previous estimation and updating the seat allocation plan should be completed in a timely manner to avoid the perception by the customer of a long delay.

The company must actively and continually revise the previous estimation to update the allocation plan. In our opinion, that is the essence of the SR heuristic. Without revising the estimation, we can not apply Theorem 3 to justify the optimality of SR heuristic. We will show that simply resolving the stochastic program without revising previous estimation will not necessarily benefit the booking process.

Example 1: Consider a single leg booking with two booking classes, business and leisure and 10 seats. The fares are \$300 and \$100 for business and leisure customers respectively. The demands for both classes are $[4; 8]$ with probability one. If we divide the booking horizon into two sub-periods, the underlying distribution for customer demands is $[2; 4]$ with probability one for each sub-period. Thus, the optimal solution will be to allocate 4 seats for business customers and the remaining 6 seats for leisure customers, with the total revenue at

$$4 \times \$300 + 6 \times \$100 = \$1800 \tag{3.5}$$

which is the revenue without re-solving. However, the company encounters incomplete distributional information. For the first sub-period, the realised demands are $[3; 4]$ and the second sub-period has a demand of $[1; 4]$. When we re-solve the model at the starting point of the second sub-period, we have 3 seats left with an underlying distribution of $[2; 4]$ with probability one. Thus, the *revised* allocation becomes $[2; 1]$ and thereby the total revenue becomes

$$3 \times \$300 + 4 \times \$100 + 1 \times \$300 + 1 \times \$100 = \$1700 \tag{3.6}$$

Comparing to the non-resolving method, the re-solving method costs an additional \$100. For a mid-scale problem, we present the numerical evidence in Section 4 to show that the re-solving method, without accurate distributional information, indeed backfires the model performance.

3.2 Risk-averse objective

In general, the stochastic programming model for the airline revenue management is risk-neutral (see Talluri and Van Ryzin (2005)) because of the scale of operation and the volume of daily demands. For some quantity-based revenue management systems, such as car rental and hotel, the companies may have their own risk preferences that are overwhelmingly risk-averse. The customers in these industries can still belong to many booking classes which pay fares at different levels. The state-of-the-art research on the risk-averse seat allocation models is in Levin et al. (2008). In this research, we adopt the MV objective as follows:

$$\begin{aligned} &\max f'x + \mathbb{E}[Q(x,\xi)] - \frac{\gamma}{2}\mathrm{Var}[Q(x,\xi)] \\ \text{subject to:} \quad & Ax \leq c \\ & x \in \mathbb{R}^r_+ \end{aligned} \tag{SLP-MV}$$

where $\gamma > 0$ is a parameter that indicates risk attitude. We choose the MV risk-averse objective for two major practical benefits. First, the MV objective is implementable because it requires only two moments that can be easily estimated. Second, this model is useful in the sense that the MV objective corresponds to a CARA utility function with normal noise. Regardless, the company may although very unlikely for revenue management companies use a non-monetary utility (see Van Mieghem (2007)).

The MV objective is only piecewise concave; it is not concave with respect to x (see Ahmed (2006)). Consequently, equation (SLP-MV) is not a convex optimisation model, and the optimal solution obtained might be only local optimal. Nevertheless, when the following condition is satisfied, the MV objective in equation (SLP-MV) is indeed concave. (For technical detail, we refer readers to Ruszczynski and Shapiro (2003), Chapter 1).

Assumption 1: The distribution of $Q(x, \xi)$ is symmetrical around its mean, i.e., $\mathbb{E}[Q(x, \xi)]$.

If the Assumption 1 is satisfied for every x such that $Ax \leq c$, model (SLP-MV) has a concave objective and its optimal solution is global. Furthermore, we pose the following proposition:

Proposition 5: *When the company becomes more risk-averse, the optimal solution and optimal value decrease with respect to an increasing γ.*

Proof: Given $\forall\, \gamma_1,\, \gamma_2$, such that $\gamma_2 > \gamma_1 \geq 0$. Let $U(x; \gamma_1)$ and $U(x; \gamma_2)$ be the objectives of equation (SLP-MV) with parameter γ_1 and γ_2, respectively. As previously discussed, $\mathbb{E}[Q(x, \xi)]$ and $\mathbb{E}[Q(x, \tilde{\xi})]$ are non-differentiable, we need the sub-gradient argument which is mainly about the set operation. Without loss of generality, we assume $\mathbb{E}[Q(x, \xi)]$ and $\mathbb{E}[Q(x, \tilde{\xi})]$ differentiable for the sake of concise writing. Thus,

$$\frac{\partial U(x; \gamma_1)}{\partial x} - \frac{\partial U(x; \gamma_2)}{\partial x} = (\gamma_2 - \gamma_1)\mathbb{E}[\text{cov}(\frac{\partial Q(x, \xi)}{\partial x}, Q(x, \xi))] \geq 0 \tag{3.7}$$

where $\mathbb{E}[\text{cov}(\frac{\partial Q(x, \xi)}{\partial x}, Q(x, \xi))] \geq 0$ is obtained from Theorem 236 of Hardy et al. (1964). □

This proposition demonstrates the fact that, as the company becomes more risk-averse, the optimal solution and optimal value will decrease in comparison to the risk-neutral counterparts. For a risk-averse revenue management company, we can replace the model (SLP) by equation (SLP-MV) in Steps 1 and 2 for the risk-averse version of SR heuristic. Let $\nu^*_{MV}(\mathbb{P})$ and $\nu^*_{MV}(\mathbb{Q})$ denote the optimal values; $x^*_{MV}(\mathbb{P})$ and $x^*_{MV}(\mathbb{Q})$ denote the optimal solutions of equation (SLP-MV) under distributions $\mathbb{P}$ and $\mathbb{Q}$ respectively. When we assume that the sequence of gradually improving estimations $\{\mathbb{P}_{t_\ell}\}$ is weakly convergent to the real distribution $\mathbb{Q}$, by applying Theorem 3, the sequence of optimal values of model (SLP-MV) will converge to the true optimal solution, $\nu^*_{MV}(\mathbb{Q})$. The sequence of optimal solutions of model (SLP-MV) will become closer to the true optimal solutions of risk-averse model as well. We still use $\theta \in [0, 1]$ to purposefully underestimate demands. In the phase I, we continue booking when the following condition is satisfied.

$$y_{t_\ell} \leq \min\{x^*_{MV}(\mathbb{P}_{t_{\ell-1}}), x^*_{MV,\theta}(\mathbb{P}_{t_{\ell-1}})\}, t_\ell \in [0, t_l) \tag{3.7}$$

where $x^*_{MV,\theta}(\mathbb{P}_{t_{\ell-1}}) :=$ argmin$\{f'x + \mathbb{E}[Q(x, \theta\xi)] - \frac{\gamma}{2}\text{Var}[Q(x, \theta\xi)] | Ax \leq c, x \geq 0\}$. Otherwise, we need to update the estimates and test the following condition

$$y_{t_\ell} \leq \min\{x^*_{MV}(\mathbb{P}_{t_\ell}), x^*_{MV,\theta}(\mathbb{P}_{t_\ell})\}, t_\ell \in [0, t_l) \tag{3.7}$$

If the rejection is confirmed, then the phase I is over. In phase II, we continue booking under the condition

$$(y_{t_\ell} - z_{t_\ell}) \cdot H \leq \min\{x^*_{MV}(\mathbb{P}_{t_{\ell-1}}), x^*_{MV,\theta}(\mathbb{P}_{t_{\ell-1}})\} \cdot H, t_\ell \in [t_l, \tau] \tag{3.7}$$

If any high-fare rejection is about to happen, we update the estimate to either confirm the rejection or adjust the allocation plan. We adjust the allocation to avoid rejecting high-fare booking requests if the following is satisfied.

$$(y_{t_\ell} - z_{t_\ell}) \cdot H \leq \min\{x^*_{MV}(\mathbb{P}_{t_\ell}), x^*_{MV,\theta}(\mathbb{P}_{t_\ell})\} \cdot H, t_\ell \in [t_l, \tau]. \tag{3.8}$$

If equation (3.8) fails, we confirm the high-fare rejection for certain booking class. In addition, for risk-averse models, when overestimation (underestimation) occurs, equation (SLP-MV) tends to allocate more (fewer) seats in comparison to the true optimal solution. This tendency for the risk-averse model is consistent with the result of the risk-neutral model because those results are developed from generic stochastic programs with convex objectives regardless the risk measure adopted.

4 Numerical study

In this section we describe the results from numerical experiments performed with the SR heuristic. Although our dataset was randomly generated, we tried to mimic real data as much as possible. To do so, we imposed the following features, which according to Belobaba and Weatherford (1996) are characteristic of actual booking processes. They are

- uncertain but limited number of potential customers
- uncertain mix of high or low-fare customers
- uncertain order of arrivals
- high-fare customers tend to arrive after the low-fare ones.

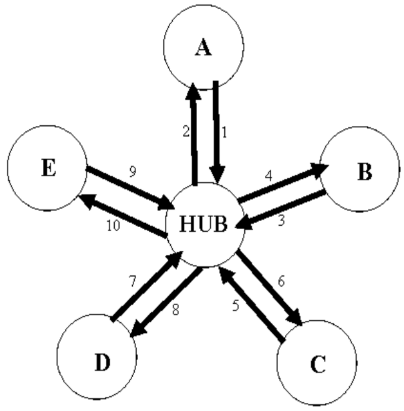
We use the same examples in Chen and Homem-de-Mello (2010b), and of course, we demonstrate that with incomplete distributional information, the proposed SR heuristic will numerically outperform traditional stochastic programs.

Example 2: This example is a 10-leg network described in Figure 1. This network is similar to the spoke-and-hub network structure in practical operations for major airlines. We consider all flights to/from the hub from/to each city, as well as the flight between two cities connecting at the hub. Therefore, there are 30 possible itineraries in the network. There are two booking classes for each flight, with the proportion of 1:3 between high and low-fare classes in terms of total requests. Following Liu and Van Ryzin (2011), we model the booking process of the *underlying distribution* $\mathbb{P}_i$ by a doubly stochastic non-homogeneous Poisson process, where the arrival intensity at time t has a gamma distribution. More specifically, for each itinerary j let $\lambda_{j1}(t)$ and $\lambda_{j2}(t)$ be the arrival intensity of respectively high-fare and low-fare customers at time t. Denote by $\alpha_j > 0$ the expected total number of requests for itinerary j over the booking horizon (i.e., for both classes together). Let G_j be a random variable with gamma distribution with shape parameter α_j and scale parameter $\beta' = 1$ (That is, the density function of G_j is $f_j(x) = \dfrac{(x/\beta')^{\alpha_j - 1} d^{-x/\beta'}}{\beta'\Gamma(\alpha_j)}, x \geq 0$).

We define $\lambda_{jk}(t), k = 1, 2$ as

$$\lambda_{jk}(t) = \beta_{jk}(t) \times G_j \times \psi_k \tag{4.1}$$

Figure 1 Example 2



where

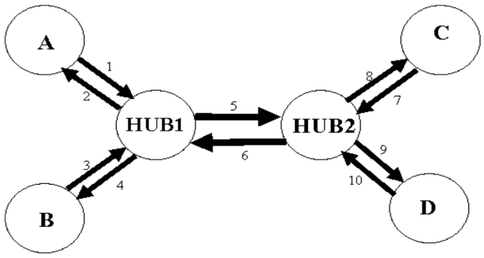
$$\beta_{jk}(t) = \frac{1}{\tau}\left(\frac{t}{\tau}\right)^{a_{jk}-1}\left(1-\frac{t}{\tau}\right)^{b_{jk}-1}\frac{\Gamma(a_{jk}+b_{jk})}{\Gamma(a_{jk})\Gamma(b_{jk})} \tag{4.2}$$

The parameters ψ_1, ψ_2 are set with the goal of reflecting the proportion of arrivals for high and low-fare customers. We take this proportion to be 1:3 in all the itineraries, so we set $\psi_1 = 0.25$ and $\psi_2 = 0.75$. Notice that for each t, $\lambda_{jk}(t)$ has gamma distribution with shape parameter α_j and scale parameter $\beta_{jk}(t)\psi_k$. In particular, $\mathbb{E}[\lambda_{jk}] = \alpha_j\beta_{jk}(t)\psi_k$, and hence the total expected number of arrivals for itinerary j is $\int_0^\tau \mathbb{E}[\lambda_{j1}(t)] + \mathbb{E}[\lambda_{j2}(t)]dt = \alpha_j\psi_1 + \alpha_j\psi_2 = \alpha_j$, which is consistent with our definition of α_j.

The parameters $\beta_{jk}(t)$ are selected to reflect the arrival patterns of different classes. High-fare customers tend to arrive close to the end of the booking horizon, whereas low-fare customers usually appear early in the booking process. To model that, we set $a_{j1} > b_{j1}$, and $a_{j2} < b_{j2}$. In our example we used $a_{jk}, b_{jk} \in \{2, 6\}$ for all j, k. From Figure 1, we see that there are 10 one-leg and 20 two-leg itineraries. For two-leg itineraries, we set the total expected number of requests equal to 100, that is $a_j = 100$. The high and low fares are $f_{j1} = \$500$ and $f_{j2} = \$100$, respectively. For one-leg itineraries, we set the *underlying* expected total number of requests equal to 40, that is $\alpha_j = 40$, with the high and low fares set as $f_{j1} = \$300$ and $f_{j2} = \$80$, respectively. All legs in the network have capacity equal to 400, and the booking horizon has a length of $\tau = 1,000$ time units.

Example 3: This example is depicted in Figure 2, which reflects a typical class of itinerary for passengers in practice. Again, we consider two classes for each itinerary. Notice that there are 10 one-leg, 12 two-leg, and 8 three-leg itineraries. The expected number of requests is 60, 150, and 100, respectively. The fare levels for different types of itineraries are set as $(\$300, \$80)$, $(\$500, \$100)$, and $(\$700, \$200)$. The parameters a_{jk}, b_{jk}, ψ_1 and ψ_2 as well as the horizon length are the same as in the first example. The leg capacities are 400 for the arcs connecting the hubs with the satellite nodes and 1,000 for the arcs connecting the two hubs.

Figure 2 Example 3



The most important part of these two numerical experiments is the performance of the SR heuristic with incomplete distributional information. In these experiments, we model the customer arrival as the Poisson process with *incomplete distributional information* as follows.

Case A. Changing the nominal $\lambda_{j1}(t), \lambda_{j2}(t)$ by *randomly* perturbing $\pm 25\%$.

Case B. Changing the nominal $\lambda_{j1}(t), \lambda_{j2}(t)$ by *randomly* perturbing $\pm 50\%$.

Case C. Unexpected high demand levels $\lambda_{j1}(t), \lambda_{j2}(t)$ by 300% on certain booking classes.

We denote the realised arrival rates by $\hat{\lambda}_{j1}(t)$ and $\hat{\lambda}_{j1}(t)$.

The SR heuristic will partially allocate seats and learn the realised distributional information simultaneously. In the numerical experiment, we define a parameter $1 \geq \beta(t) > 0, t = 0, \ldots, \tau$ as a measure of the gradual *learning* on the realised distribution $\mathbb{Q}$ such that $\beta(0) = 0, \beta(\tau) = 1$. During the simulation, we use $\lambda_{ij}(t) + \beta(t)[\hat{\lambda}_{ij}(t) - \lambda_{ij}(t)]$ in models (SLP) and (SLP-MV) for the *revised* allocation plan. We model the *learning* by

$$\beta(t) = \begin{cases} \beta(0) + \dfrac{1-\beta(0)}{800} t & t \in [1, 800] \\ 1 & \text{Otherwise} \end{cases} \tag{4.3}$$

which means that our knowledge on the perturbation increases linearly and for $t \geq 801$, our estimation coincides with the realised distribution $\mathbb{Q}$, i.e., $\beta(t) = 1$. We model the company's conservative seat allocation plan by fixing θ throughout the entire booking horizon. $\theta < 1$ means that the company is *purposefully underestimating* the demand. We present the summaries in Tables 1 and 2 with the following abbreviations: RSP-2–resolving stochastic models twice at time points 1 and 501; RSP-5–resolving stochastic models five times at time points $1, 201, 401, 601, 801$; SRI–Improvements in terms of expected total revenue by SR heuristic; SR – seat reservation.

In both examples, the SR heuristic outperforms the re-solving stochastic models and the improvement could be as good as 9% (when $1.5\lambda_{j1}(t), 1.5\lambda_{j2}(t)$). Considering the scale of airline daily operation, the resulting savings on revenue should be in the scale of millions of dollars. In particular, when the high-fare customer demands are strong, the SR heuristic will yield more revenue because it dynamically releases seats to revisit the previously made allocations with more information revealed, thereby better estimations. We use the COIN-OR (see http://www.coin-or.org) solver.

Table 1 Numerical summary for Example 2 with 1000 replications

Case settings	*RSP-2*	*RSP-5*	*SR*, $\theta = 0.80$	*SRI* (%)
$1.25\lambda_{j1}(t), \lambda_{j2}(t)$	450700 ± 407	456340 ± 353	464500 ± 560	1.78
$\lambda_{j1}(t), 1.25\lambda_{j2}(t)$	422210 ± 582	427300 ± 620	428391 ± 594	0.2
$1.5\lambda_{j1}(t), \lambda_{j2}(t)$	463070 ± 243	467420 ± 165	489820 ± 305	4.7
$\lambda_{j1}(t), 1.5\lambda_{j2}(t)$	425190 ± 581	434730 ± 562	436000 ± 540	0.2
$1.25\lambda_{j1}(t), 1.25\lambda_{j2}(t)$	455490 ± 364	460130 ± 300	458520 ± 348	-0.3
$1.5\lambda_{j1}(t), 1.5\lambda_{j2}(t)$	469210 ± 187	470350 ± 129	503970 ± 295	8.9
$0.75\lambda_{j1}(t), \lambda_{j2}(t)$	349220 ± 679	351180 ± 665	349550 ± 672	-0.4
$\lambda_{j1}(t), 0.75\lambda_{j2}(t)$	407450 ± 598	411140 ± 657	409730 ± 647	-0.3
$0.5\lambda_{j1}(t), \lambda_{j2}(t)$	252740 ± 578	254170 ± 549	252770 ± 540	-0.5
$\lambda_{j1}(t), 0.5\lambda_{j2}(t)$	397320 ± 580	401760 ± 652	401810 ± 667	0.01
$0.75\lambda_{j1}(t), 0.75\lambda_{j2}(t)$	338970 ± 658	339790 ± 655	339960 ± 688	0.05
$0.5\lambda_{j1}(t), 0.5\lambda_{j2}(t)$	227120 ± 554	226480 ± 549	226790 ± 545	-0.14
$0.75\lambda_{j1}(t), 1.25\lambda_{j2}(t)$	358160 ± 625	362510 ± 648	362180 ± 587	-0.09
$0.5\lambda_{j1}(t), 1.5\lambda_{j2}(t)$	269180 ± 556	279040 ± 582	278950 ± 547	-0.03
$1.25\lambda_{j1}(t), 0.75\lambda_{j2}(t)$	444080 ± 415	451310 ± 399	475420 ± 358	5.3
$1.5\lambda_{j1}(t), 0.5\lambda_{j2}(t)$	452450 ± 250	462720 ± 223	491350 ± 214	6.2
$3\lambda_{j1}(t), \lambda_{j2}(t)$	471100 ± 177	468820 ± 132	691550 ± 889	46.8
$\lambda_{j1}(t), 3\lambda_{j2}(t)$	430560 ± 234	452480 ± 562	471800 ± 796	4.2

Table 2 Numerical summary for Example 3 with 1000 replications

Case settings	*RSP-2*	*RSP-5*	*SR*, $\theta = 0.80$	*SRI* (%)
$1.25\lambda_{j1}(t), \lambda_{j2}(t)$	626870 ± 597	636450 ± 510	667440 ± 498	4.8
$\lambda_{j1}(t), 1.25\lambda_{j2}(t)$	599320 ± 576	601890 ± 563	603300 ± 777	0.2
$1.5\lambda_{j1}(t), \lambda_{j2}(t)$	661630 ± 528	662320 ± 368	669610 ± 452	1.1
$\lambda_{j1}(t), 1.5\lambda_{j2}(t)$	599820 ± 517	604770 ± 433	602490 ± 470	-0.3
$1.25\lambda_{j1}(t), 1.25\lambda_{j2}(t)$	645210 ± 369	649990 ± 309	665620 ± 263	2.4
$1.5\lambda_{j1}(t), 1.5\lambda_{j2}(t)$	659480 ± 169	662140 ± 190	713950 ± 895	7.8
$0.75\lambda_{j1}(t), \lambda_{j2}(t)$	501210 ± 807	500100 ± 789	499150 ± 312	-0.4
$\lambda_{j1}(t), 0.75\lambda_{j2}(t)$	580950 ± 816	585050 ± 954	581120 ± 599	-0.6
$0.5\lambda_{j1}(t), \lambda_{j2}(t)$	407210 ± 732	406240 ± 719	399780 ± 176	-1.5
$\lambda_{j1}(t), 0.5\lambda_{j2}(t)$	522540 ± 727	538570 ± 701	543850 ± 865	0.9
$0.75\lambda_{j1}(t), 0.75\lambda_{j2}(t)$	485340 ± 810	478690 ± 736	490950 ± 566	1.2
$0.5\lambda_{j1}(t), 0.5\lambda_{j2}(t)$	320370 ± 699	336140 ± 897	348850 ± 838	3.7
$0.75\lambda_{j1}(t), 1.25\lambda_{j2}(t)$	510960 ± 678	511790 ± 620	505140 ± 305	-1.2
$0.5\lambda_{j1}(t), 1.5\lambda_{j2}(t)$	406670 ± 592	405880 ± 503	397650 ± 152	-2.2
$1.25\lambda_{j1}(t), 0.75\lambda_{j2}(t)$	650080 ± 378	654700 ± 332	655660 ± 893	0.2
$1.5\lambda_{j1}(t), 0.5\lambda_{j2}(t)$	671040 ± 779	687330 ± 198	731760 ± 731	6.4
$3\lambda_{j1}(t), \lambda_{j2}(t)$	673130 ± 253	671570 ± 193	975860 ± 911	44.9
$\lambda_{j1}(t), 3\lambda_{j2}(t)$	599150 ± 550	603870 ± 578	606080 ± 450	0.4

We summarise our observations on numerical experiment into the following points: first, the SR heuristic with better estimations will not necessarily lead to a greater expected revenue. There are small negative SRI observed. However, these negative SRI values are reasonably tight and nearly negligible. Secondly, the SR heuristic will be extremely important under the large perturbations of the high-fare classes, e.g., $\hat{\lambda}_{j1}(t) = 3\lambda_{j1}(t)$. The resulting revenue jump observed is up to 45% compared with its counterpart of simply re-solving stochastic programs without updating the demand estimation. Similar results about stochastic seat allocation system can be found in Maglaras and Meissner (2006).

5 Conclusions

We have discussed the airline booking process based on the origin-destination model with incomplete distributional information. More specifically, we have presented a heuristic, the SR heuristic, to reduce the risk whenever wrong estimations or erroneous information are encountered. Our proposed SR heuristic is to adaptively incorporate available information into the seat allocation plan. By continually revising the previous estimations and incorporate more available information, we formulate a sequence of distributions which is assumed to be weakly convergent to the true distribution. Moreover, we found that the revenue management generated from SR method *could* coincides with the true optimal value. Thus, by applying Theorem 3, we conclude that the SR heuristic will provide a satisfactory solution that can be arbitrarily close to the true optimal value. Beyond the theoretical analysis, the proposed SR heuristic not only improves the model performance in terms of total expected revenue but also provides the insightful thoughts. In particular, the SR heuristic suggests three important components for the revenue management company: the conservative decision, the active and continual revisions on estimations, and an adequately up-to-date computational facility.

Our analysis suggests that the proposed method is robust in the sense that solving successive stochastic models with updated estimations can not only improve the expected revenue, but also greatly reduce the risk associated with incomplete distributional information. We discuss the SR heuristic both for risk-neutral and risk-averse companies by the well-known MV risk objective. The consistency between the results from risk-neutral and risk-averse models indicates that our SR heuristic can be extended to other quantity-based revenue management systems with different risk attitude. In addition, we show that both of the revenue management models are convex optimisation models under mild assumptions. Thus, the SR heuristic for both risk-neutral and risk-averse companies, which formulates a sequence of convex optimisations, can be solved efficiently in a timely manner, both practically and theoretically. For example, with an adequately up-to-date computational facility, the customers who place requests during the events of revising estimation and updates allocation plan, are less likely to perceive such a delay caused by computation.

We must remark that we have not included in our models some of the recent developments proposed in the literature, e.g., nesting of classes and the customer choice model. There are two basic reasons for our decision. First, incorporation of the above features leads to different models that lie outside of the scope of this paper due to the potential loss of convexity by the simulation based methods (see Bertsimas and de Boer (2005)) and the customer choice models (see Chen and Homem-de-Mello (2010a)). Second, our conversations with people in the airline industry have shown us that the basic origin-destination model, particularly the deterministic linear programming formulation, is widely used in practice with the nesting operation. Thus, our goal is to provide the practitioners an easily implementable algorithm that can improve the booking process in terms of lower risk, i.e., variation and greater expected revenue. Again, our theoretical discussion and numerical results show that the SR heuristic proposed in this paper will soundly accomplish the goal.

References

Ahmed, S. (2006) 'Convexity and decomposition of mean-risk stochastic programs', *Mathematical Programming*, Vol. 106, No. 3, pp.433–446.

Aydin, N., Birbil, S.I., Frenk, J. and Noyan, N. (2012) 'Single-leg airline revenue management with overbooking', *Transportation Science*, Vol. 47, No. 4, pp.560–583.

Belobaba, P. and Weatherford, L. (1996) 'Comparing decision rules that incorporate customer diversion in perishable asset revenue management situations', *Decision Sciences*, Vol. 27, No. 2, pp.343–363.

Ben-Tal, A. and Nemirovski, A. (1999) 'Robust solutions of uncertain linear programs', *Operations Research Letters*, Vol. 25, No. 1, pp.1–14.

Bertsimas, D. and de Boer, S. (2005) 'Simulation-based booking limits for airline revenue management', *Operations Research* pp.90–106.

Chen, L. and Homem-de-Mello, T. (2010a) 'Mathematical programming models for revenue management under customer choice', *European Journal of Operational Research*, Vol. 203, No. 2, pp.294–305.

Chen, L. and Homem-de-Mello, T. (2010b) 'Re-solving stochastic programming models for airline revenue management', *Annals of Operations Research*, Vol. 177, No. 1, pp.91–114.

Cooper, W. (2002) 'Asymptotic behavior of an allocation policy for revenue management', *Operations Research,* pp.720–727.

Ferrer, J-C., Oyarzún, D. and Vera, J. (2012) 'Risk averse retail pricing with robust demand forecasting', *International Journal of Production Economics*, Vol. 136, No. 1, pp.151–160.

Graf, M. and Kimms, A. (2011) 'An option-based revenue management procedure for strategic airline alliances', *European Journal of Operational Research*, Vol. 215, No. 2, pp.459–469.

Hardy, G., Littlewood, J. and Polya, G. (1964) *Inequalities*, 2nd ed., Cambridge University Press, Cambridge, UK.

Huang, K. and Chang, K-C. (2011) 'A model for airline seat control considering revenue uncertainty and risk', *Journal of Revenue and Pricing Management*, Vol. 10, No. 2, pp.161–171.

Lemke, C., Riedel, S. and Gabrys, B. (2013) 'Evolving forecast combination structures for airline revenue management', *Journal of Revenue and Pricing Management*, Vol. 12, No. 3, pp.221–234.

Levin, Y., McGill, J. and Nediak, M. (2008) 'Risk in revenue management and dynamic pricing', *Operations Research*, Vol. 56, No. 2, p.326.

Liu, Q. and Van Ryzin, G. (2011) 'Strategic capacity rationing when customers learn', *Manufacturing and Service Operations Management*, Vol. 13, No. 1, pp.89–107.

Maglaras, C. and Meissner, J. (2006) 'Dynamic pricing strategies for multi-product revenue management problems', *Manufacturing and Service Operations Management*, Vol. 8, No. 2, pp.136–148.

Nemirovski, A., Juditsky, A., Lan, G. and Shapiro, A. (2009), 'Robust stochastic approximation approach to stochastic programming', *SIAM Journal on Optimization*, Vol. 19, No. 4, pp.1574–1609.

Rachev, S. and Römisch, W. (2002) 'Quantitative stability in stochastic programming: the method of probability metrics', *Mathematics of Operations Research*, pp.792–818.

Rockafellar, R. and Wets, R.J. (1998) *Variational Analysis*, Springer Verlag, Berlin, Germany, Vol. 317.

Romisch, W. and Schultz, R. (1993) 'Stability of solutions for stochastic programs with complete recourse', *Mathematics of Operations Research*, Vol. 18, No. 3, pp.590–609.

Ruszczynski, A. and Shapiro, A. (Eds.) (2003) *Stochastic Programming*, Vol.10 of *Handbooks in Operations Research and Management Science*, Elsevier, New York, NY.

Talluri, K. and Van Ryzin, G. (1998) 'An analysis of bid-price controls for network revenue management', *Management Science*, pp.1577–1593.

Talluri, K. and Van Ryzin, G. (2005), *The Theory and Practice of Revenue Management*, Springer Verlag, Berlin, Germany, Vol. 68.

Van Mieghem, J. (2007) 'Risk mitigation in newsvendor networks: resource diversification, flexibility, sharing, and hedging', *Management Science*, Vol. 53, No. 8, p.1269.

Zhuang, W. and Li, M.Z. (2011) 'Revenue optimization of risk-averse managers with atemporal utility', *Journal of Revenue and Pricing Management*, Vol. 10, No. 5, pp.424–437.