Submitted to *Management Science*
manuscript

The Role of Robust Optimization in Single-leg Airline Revenue Management

Ş. İlker Birbil
Faculty of Engineering and Natural Sciences, Sabancı University, 34956 Istanbul, Turkey, sibirbil@sabanciuniv.edu, http://people.sabanciuniv.edu/sibirbil

J.B.G. Frenk
Erasmus University Rotterdam, Postbus 1738, 3000 DR Rotterdam, The Netherlands, frenk@few.eur.nl, http://people.few.eur.nl/frenk/

Joaquim A.S. Gromicho
Vrije Universiteit, Amsterdam & ORTEC, Gouda, The Netherlands, jgromicho@feweb.vu.nl, http://www.ortec.nl

Shuzhong Zhang
Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, HK, zhang@se.cuhk.edu.hkl, http://www.se.cuhk.edu.hk/ zhang

In this paper we introduce robust versions of the classical static and dynamic single leg seat allocation models as analyzed by Wollmer, and Lautenbacher and Stidham, respectively. These robust models take into account the inaccurate estimates of the underlying probability distributions. As observed by simulation experiments it turns out that for these robust versions the variability compared to their classical counterparts is considerably reduced with a negligible decrease in average revenue.

Key words: airline revenue management, single-leg problems, static models, dynamic models, robust optimization

History: –

1. Introduction

Airline seat allocation problems on single legs or networks play a prominent role within the revenue management literature. This field expanded rapidly in recent years and for an overview on revenue management up to 1999 we refer the reader to McGill, J.I., G.J. Van Ryzin (1999), while developments occurring after this work are discussed in the recent book by Talluri, K.T., G.J. Van Ryzin (2004). Although many practical seat allocation problems observed in the airline industry are network based, single leg seat allocation problems still play an important role. This is mainly due to two reasons: Firstly, in general the network based airline seat allocation problems are extremely difficult to solve. Therefore, different heuristics, which require the solution of many single leg problems, were developed. Secondly, some small airline companies, like charter flight companies commonly seen in Europe, have special one-hub networks with single legs. Therefore for those companies managing their seat allocation over the network requires solving only single leg problems. Among the single leg problems, one may distinguish static and dynamic models. Further, the static models can be categorized into two types. The first type assumes that only the distribution of the demand for different fare classes is known. Since the objective is to maximize the expected revenue, this leads to the formulation of mathematical programming models. Examples of such models are given in Wollmer, R.D. (1986), De Boer, S.V., R. Freling, N. Piersma (2002), Talluri, K.T., G.J. Van Ryzin (2004). The second type assumes that the demands for different fare classes arrive in non overlapping time periods in the order of increasing fare class prices. Given a realization of a particular fare class demand, one needs to decide how much of this demand is allocated to seats, under the probabilistic information on the demand for the remaining higher priced fare classes. This model can be solved by dynamic programming, where the stages correspond to fare classes. Examples of such models under different assumptions are presented for two fare classes in Littlewood, K. (1972), Richter, H. (1982), and for more than two fare classes in Belobaba, P.P. (1989) (a heuristic approach generalizing the rule of Littlewood) and also in Wollmer, R.D. (1992), Brumelle, S.L, J.I. McGill (1993), Robinson, L.W. (1995). Finally, dynamic single leg models take into account the actual order of arrival of different fare class customers and so the decision to accept or reject a specific fare class customer is not static, but may change over time. In this case stages correspond to time periods. Examples of such models under different assumptions are given in Lee, T.C., M. Hersch (1993), Lautenbacher, C.J., S. Stidham Jr. (1999), Walczak, D., S. Brumelle (2003).

In this paper we first review, in the section on static models, the mathematical programming formulation of the static single leg problem already given by Wollmer, R.D. (1986) in a more complicated network environment (see also De Boer, S.V., R. Freling, N. Piersma (2002), Talluri, K.T., G.J. Van Ryzin (2004)). However, in these references only a binary linear programming formulation is given without any special purpose algorithms to solve those formulations. For the more special single leg case considered here, we give in Section 2 a fast special purpose algorithm to solve this model. Moreover, we present also in Section 2 a new robust formulation of the mathematical programming model, which takes into account the inaccurate estimate of the probability distributions of the total demand for the different fare classes. As shown in Section 5 it will turn out in our simulation experiments that the variability of the realized revenues is considerably smaller for the robust version. At the same time due to the conservative behavior of the robust model, the average revenues for the classical single static model are slightly higher. In Section 3 we then review the standard classical dynamic single leg problem as discussed in Lautenbacher, C.J., S. Stidham Jr. (1999) and propose, also for this model, a new robust version. This robust version takes again into account the inaccurate estimates of the probabilities of the arrival process. Again from our simulation results in Section 5 we observe the same behavior as observed for the static models. In Section 4 we consider shortly which model we have to use in case of perfect information. Then we compare the three different models (static, dynamic and perfect information) extensively by means of simulation in Section 5. Our simulation results show that the cost of lacking perfect information is relatively small. Finally, in Section 6 we conclude the paper.

2. Static Models

In this section we are interested in the optimal allocation of the seat capacity C on a given flight among the m different fare classes. If the demand d_i for each fare class i, $1 \le i \le m$, is known in advance, it is trivial to solve this allocation problem which can be modeled in the following way. Let x_i denote the number of reserved seats for fare class i at the beginning of the booking period. We assume that fare class i customers do not consider the possibility of buying a ticket from a different fare class. Thus, once no fare class i ticket is available, then it follows that $\min\{x_i, d_i\}$ will be the number of occupied fare class i seats on the selected flight. Let r_i denote the price of a fare class i seat and assume without loss of generality that $r_1 < r_2 < ... < r_m$. Then, to determine the optimal allocation of the different fare classes over the given capacity C, we need to solve the following optimization problem

$$\begin{aligned} v_1(C) := \max\ & \textstyle\sum_{i=1}^m r_i \min\{x_i, d_i\} \\ \text{s.t.}\ & \textstyle\sum_{i=1}^m x_i \le C, \\ & x \in \mathbb{Z}_+^m. \end{aligned} \tag{1}$$

It is obvious that an optimal allocation is given as follows. Consider demand d_i and price r_i for each fare class i, and assign all the seats to the higher-priced customers as long as the capacity is still available. To formalize the algorithm, introduce $S_n := \sum_{j=n}^m d_j$ with $d_0 := 0$ and $N(C) = \min\{0 \le n \le m \mid S_n \le C\}$. Then, the optimal solution of optimization problem (1) is given by

$$x_i^* = \begin{cases} d_i, & \text{if } i \ge N(C) \\ C - S_{N(C)}, & \text{if } i = N(C) - 1 \\ 0, & \text{if } i < N(C) - 1. \end{cases} \tag{2}$$

The associated optimal objective function value as a function of the capacity C is given by

$$v_1(C) = \sum_{i=N(C)}^m r_i d_i + \left(C - S_{N(C)}\right) r_{N(C)-1},$$

which is, clearly, a piecewise linear concave function.

However, usually the demand for fare class i is a random variable $\mathbf{D}_i$ and we do not know in advance its realization. We may, however, estimate the distribution of the demand. Let $\mathbf{D}_i(\omega)$ be a realization of the demand $\mathbf{D}_i$ and x_i be the number of reserved seats for fare class i. Consequently, the total revenue is given by $\sum_{i=1}^m r_i \min\{x_i, \mathbf{D}_i(\omega)\}$. This shows that the expected revenue equals $\sum_{i=1}^m r_i \mathbb{E}\left[\min\{x_i, \mathbf{D}_i\}\right]$, and so, our *static* decision model for random demand is given by

$$\begin{aligned} v_2(C) := \max\ & \textstyle\sum_{i=1}^m r_i \mathbb{E}\left[\min\{x_i, \mathbf{D}_i\}\right] \\ \text{s.t.}\ & \textstyle\sum_{i=1}^m x_i \le C, \\ & x \in \mathbb{Z}_+^m. \end{aligned} \tag{3}$$

This static model was first formulated by Wollmer, R.D. (1986) in a much more complicated network environment and became a classical model in airline seat management. Since the simpler single-leg version is a standard separable problem, it can be solved by dynamic programming, where the fare classes and the airline capacity correspond to the stages and the state space, respectively. Introduce for every $p \le m$ and $y \in \{0, ..., C\}$ the value $R_p(y)$ as the maximal expected revenue for fare classes p up to m if at most capacity y is reserved for those fare classes, i.e.,

$$R_p(y) = \max\left\{ \sum_{i=p}^m r_i \mathbb{E}\left[\min\{x_i, \mathbf{D}_i\}\right] \mid \sum_{i=p}^m x_i \le y, x_i \in \mathbb{Z}, i = p, \cdots, m \right\}.$$

By the optimality principle of Bellman it now follows for every $y \in \{0, \cdots, C\}$ and $p+1 \le m$ that

$$R_p(y) = \max_{0 \le x_p \le y} \{R_{p+1}(y - x_p) + r_p \mathbb{E}[\min\{x_p, \mathbf{D}_p\}]\}.$$

Since clearly $R_m(y) = r_m \mathbb{E}[\min\{y, \mathbf{D}_m\}]$, $y \in \{0, 1, ..., C\}$, we can recursively compute the optimal objective value $R_1(C)$. The computational complexity of this dynamic programming approach is of the order of $O(mC^2)$.

Clearly, to apply this approach we need an efficient algorithm to compute the function values $\mathbb{E}[\min\{x_i, \mathbf{D}_i\}]$. This can be done in a direct way for some simple distributions or in case the generating function of the demand has a nice analytical form using the so-called Fast Fourier Transform (FFT) approach (cf. Golub, G.H., C.F. Van Loan (1996)).

2.1. An Improved Algorithm

The key idea behind our approach is to rewrite the separable objective function of problem (3). We introduce the function $F_i : \mathbb{Z} \to R$ given by

$$F_i(n) := \mathbb{E}[\min\{n, \mathbf{D}_i\}] \tag{4}$$

and observe for given $n \in \mathbb{Z}_+$ that

$$F_i(n) = \sum_{j=1}^{n} \mathbb{P}\{\mathbf{D}_i \ge j\}.$$

Using this, it is obvious that F_i is a discrete concave function; i.e., the difference $F_i(n) - F_i(n-1)$ is non-increasing in n. By relation (4), problem (3) can be rewritten as

$$\begin{array}{rl} v_2(C) = \max & \sum_{i=1}^{m} r_i F_i(x_i) \\ \text{s.t.} & \sum_{i=1}^{m} x_i \le C, \\ & x \in \mathbb{Z}_+^m. \end{array}$$

Clearly, $x_i \le C$ in this problem. Introduce now for $1 \le j \le C$, the values

$$\alpha_{ij} := F_i(j) - F_i(j-1) = \sum_{k=j}^{\infty} p_{ik},$$

where $p_{ik} = \mathbb{P}\{\mathbf{D}_i = k\}$. Notice that the objective function is separable. Therefore, $r_i \alpha_{ij}$ gives the *marginal* value of increasing x_i from $j-1$ to j. After this observation, we can solve problem (3) very fast. To explain the algorithm, we first introduce the following $m \times C$ matrix

$$\begin{bmatrix} r_1\alpha_{11} & r_1\alpha_{12} & \cdots & r_1\alpha_{1C} \\ r_2\alpha_{21} & r_2\alpha_{22} & \cdots & r_2\alpha_{2C} \\ \vdots & \vdots & \cdots & \vdots \\ r_m\alpha_{m1} & r_m\alpha_{m2} & \cdots & r_m\alpha_{mC} \end{bmatrix}. \tag{5}$$

Then, the optimal objective function value $v_2(C)$ can be found by sorting the $r_i\alpha_{ij}$ values, and adding up the first C terms. Consequently, the number of times index i appears among these C terms gives the optimal solution x_i^*. Notice that since F_i is discrete concave, the marginal values in each row i are in descending order; i.e., $r_i\alpha_{i1} \geq r_i\alpha_{i2} \geq \cdots \geq r_i\alpha_{iC}$. Therefore, $v_2(C)$ can be evaluated by taking the maximum of m elements C times. The computational complexity of the proposed approach reduces to the order of $O(mC)$.

2.2. A Robust Optimization Approach

To evaluate the objective function of problem (3), we need to know the probability distribution of the customer demand. These probabilities are usually estimated by analyzing the historical data, and hence, they are prone to inaccuracies. A reasonable consideration would be: How can we immunize the model from the inaccurate data? To answer this question, we propose next a robust modeling approach.

We assume that random variable $\mathbf{D}_i$, representing the total demand for fare class i, is concentrated on $\{0, \cdots, K\}$, and this demand has an estimated probability vector $\hat{p}_i = (\hat{p}_{i0}, \cdots, \hat{p}_{iK})^\intercal$. Each $\hat{p}_{ik}$ is assumed positive. To compensate for possible estimation errors, we consider for $1 \leq i \leq m$ the probability vectors p_i belonging to the uncertainty set P_i given by

$$P_i = \{p_i \in \mathbb{R}^{K+1} \mid p_i \in \hat{p}_i + \Delta_i,\ p_i^\intercal e = 1\},$$

where

$$\Delta_i = \left\{ b_i = (b_{i0}, \cdots, b_{iK})^\intercal \in \mathbb{R}^{K+1} \mid \sum_{k=0}^{K} \left(\frac{b_{ik}}{\hat{p}_{ik}}\right)^2 \leq \delta_i^2 \right\}$$

with $\delta_i \in [0,1]$. Such type of ellipsoidal modeling for the ambiguous parameters in robust optimization was first introduced by Ben-Tal, A., A. Nemirovski (1998); see also Bertsimas, D., M. Sim (2003) for discrete problems. Recent developments of robust optimization can be found through Ben-Tal, A., L. El Ghaoui, A. Nemirovski (2006), which contains many references and useful source of information. It is easy to verify by the positivity of $\hat{p}_{ik}$ and the definition of Δ_i that $\hat{p}_i + \Delta_i \subseteq \mathbb{R}_+^{K+1}$. The total demand then depends on its probability distribution p_i, and hence we denote this random variable by $\mathbf{D}_i(p_i)$. Thus, the robust counterpart of problem (3) is given by

$$\begin{array}{rl} v_3(C) := \max & \sum_{i=1}^m r_i \min_{p_i \in P_i} \{\mathbb{E}[\min\{x_i, \mathbf{D}_i(p_i)\}]\} \\ \text{s.t.} & \sum_{i=1}^m x_i \le C, \\ & x \in \mathbb{Z}_+^m. \end{array} \tag{6}$$

We introduce then the function $G_i : \mathbb{Z}_+ \to \mathbb{R}$ given by

$$G_i(n) := \min_{p_i \in P_i} \{\mathbb{E}[\min\{n, \mathbf{D}_i(p_i)\}]\}. \tag{7}$$

Notice for every $p_i \in P_i$ that the function

$$n \to \mathbb{E}[\min\{n, \mathbf{D}_i(p_i)]$$

is discrete concave on $\mathbb{Z}_+$. Since the point-wise infimum of a collection of concave functions is again concave, the function G_i is also discrete concave on $\mathbb{Z}_+$. Then problem (6) can be rewritten as

$$\begin{array}{rl} v_3(C) = \max & \sum_{i=1}^m r_i G_i(x_i) \\ \text{s.t.} & \sum_{i=1}^m x_i \le C, \\ & x \in \mathbb{Z}_+^m. \end{array}$$

Observe for given $p_i \in P_i$ that

$$\mathbb{E}[\min\{x_i, \mathbf{D}_i(p_i)\}] = \sum_{k=0}^{x_i - 1} k p_{ik} + x_i \sum_{k=x_i}^{K} p_{ik} = c(x_i)^\intercal p_i,$$

where

$$c(x_i)^\intercal := (c_0(x_i), c_1(x_i), \cdots, c_K(x_i)) = (0, 1, \cdots, x_i - 1, x_i, x_i, \cdots, x_i).$$

Hence, by relation (7), we have

$$G_i(x_i) = \min\{c(x_i)^\intercal p_i \mid p_i \in P_i\} = c(x_i)^\intercal \hat{p}_i + \min\{c(x_i)^\intercal b_i \mid b_i \in \Delta_i,\ b_i^\intercal e = 0\}. \tag{8}$$

Using standard nonlinear programming techniques (cf. Bertsekas, D.P. (1999)), it can be easily shown that

$$\min\{c^\intercal y \mid y^\intercal Q y \le \delta^2,\ e^\intercal y = 0\} = -\delta \sqrt{c^\intercal Q^{-1} c - \frac{(e^\intercal Q^{-1} c)^2}{e^\intercal Q^{-1} e}}, \tag{9}$$

where Q is symmetric and positive definite. This shows that the last term in relation (8) has an analytic expression. Therefore, using $c_o(x_i) = 0$ we have

$$G_i(x_i) = c(x_i)^\intercal \hat{p}_i - \delta_i \sqrt{\sum_{k=1}^{K} \hat{p}_{ik}^2 c_k^2(x_i) - \frac{(\sum_{k=1}^{K} \hat{p}_{ik}^2 c_k(x_i))^2}{\sum_{k=0}^{K} \hat{p}_{ik}^2}}. \tag{10}$$

It is clear that $x_i \leq C$ in problem (6). Introduce now for $1 \leq j \leq C$, the values

$$\beta_{ij} := G_i(j) - G_i(j-1).$$

Similar to the discussion in Section 2.1, we first introduce the following $m \times C$ matrix

$$\begin{bmatrix} r_1\beta_{11} & r_1\beta_{12} & \cdots & r_1\beta_{1C} \\ r_2\beta_{21} & r_2\beta_{22} & \cdots & r_2\beta_{2C} \\ \vdots & \vdots & \cdots & \vdots \\ r_m\beta_{m1} & r_m\beta_{m2} & \cdots & r_m\beta_{mC} \end{bmatrix}. \tag{11}$$

Then, since G_i is discrete concave, the marginal values in each row i are in descending order; i.e., $r_i\beta_{i1} \geq r_i\beta_{i2} \geq \cdots \geq r_i\beta_{iC}$. Therefore, the optimal objective function value $v_3(C)$ can be evaluated by taking the maximum of m elements C times. The computational complexity of the approach to solve (6) is of the order $O(mC)$.

3. Dynamic Models

Before discussing a robust version of the dynamic single-leg problem we first review the classical dynamic single-leg problem as proposed by Lautenbacher, C.J., S. Stidham Jr. (1999). Suppose that there are m different fare classes with the prices

$$0 < r_1 < r_2 < \cdots < r_m.$$

The no-sales class is simply represented by 0 with $r_0 = 0$. The total number of available seats is denoted by z, and the ticket sales period is partitioned into periods $1, 2, \cdots, T$. We assume that in each period either no customer is observed or at most one fare class i customer arrives. If ξ_t denotes the revenue generated by this random demand in period t, we may assume that ξ_t may take $m+1$ different values $r_0, r_1, ..., r_m$ and its discrete density is given by

$$\mathbb{P}\{\xi_t = r_i\} = p_{it}$$

with $i = 0, 1, ..., m$ and $t = 1, ..., T$. It is also assumed that the random variables $\xi_t, t = 1, ..., T$ are independent. Introducing now the optimal random revenue $\mathbf{R}_t(z)$ that is generated from period t to T, before a request shows up in period t, while the number of available seats at the beginning of period t is z we denote by $J_t(z) := \mathbb{E}[\mathbf{R}_t(z)]$ the associated expected optimal value function. Clearly $J_t(z) = \mathbb{E}_{\xi_t}(\mathbb{E}[\mathbf{R}_t(z)|\xi_t])$ and by the principle of dynamic programming it follows that

$$\mathbb{E}[\mathbf{R}_t(z)|\xi_t] = \max\{\xi_t + J_{t+1}(z-1), J_{t+1}(z)\}.$$

The above equation also yields an optimal policy: Accept the request if

$$\xi_t \geq J_{t+1}(z) - J_{t+1}(z-1).$$

Therefore,

$$J_t(z) = \mathbb{E}[\max\{\xi_t + J_{t+1}(z-1), J_{t+1}(z)\}],$$

with

$$J_T(z) = \begin{cases} \mathbb{E}[\xi_T], & \text{if } z > 0 \\ 0, & \text{if } z = 0. \end{cases}$$

For the above optimal value function, the following result has been shown in Lautenbacher, C.J., S. Stidham Jr. (1999).

Theorem 1. *For every given t, the function*

$$\Delta_{t+1}(z) := J_{t+1}(z) - J_{t+1}(z-1)$$

is non-negative and non-increasing in z.

To compute the values $J_t(z)$ knowing the values $J_{t+1}(z)$ we observe

$$J_t(z) = J_{t+1}(z) + \mathbb{E}[\max\{\xi_t - \Delta_{t+1}(z), 0\}].$$

If we denote $(x)_+ = \max\{x, 0\}$, then we have

$$\mathbb{E}[\max\{\xi_t - \Delta_{t+1}(z), 0\}] = \sum_{i=0}^{m} p_{it}(r_i - \Delta_{t+1}(z))_+.$$

This yields due to $\Delta_{t+1}(z) \geq 0$ and $r_0 = 0$ that

$$J_t(z) = J_{t+1}(z) + \sum_{i=1}^{m} p_{it}(r_i - \Delta_{t+1}(z))_+. \tag{12}$$

A backward recursive solving requires an overall computational complexity of the order $O(mTC)$, where C is the total number of seats available.

3.1. A Robust Optimization Approach

In this case, the uncertain data in question are the estimated probability vectors $\hat{p}_t = (\hat{p}_{1t}, \cdots, \hat{p}_{mt})^\intercal$, $t = 1, \cdots, T$. We consider the probability vectors p_t belonging to the uncertainty set P_t given by

$$P_t = \{p_t \in \mathbb{R}^m \mid p_t \in \hat{p}_t + \Delta_t,\ p_t^\intercal e = 1\},$$

where

$$\Delta_t = \left\{ b_t = (b_{1t}, \cdots, b_{mt})^\intercal \in \mathbb{R}^m \mid \sum_{i=1}^{m} \left(\frac{b_{it}}{\hat{p}_{it}} \right)^2 \leq \delta_t^2 \right\}$$

with $\delta_t \in [0, 1]$. The dynamic programming formulation then becomes

$$J_t(z) = J_{t+1}(z) + \sum_{i=1}^{m} \hat{p}_{it}(r_i - (J_{t+1}(z) - J_{t+1}(z-1)))_+ + H_t(z)$$

with

$$H_t(z) = \min \left\{ \sum_{i=1}^{m} b_{it}(r_i - (J_{t+1}(z) - J_{t+1}(z-1)))_+ \mid b_t \in \Delta_t, e^\top b_t = 0 \right\}.$$

To simplify the notation, let

$$c_{it} := (r_i - (J_{t+1}(z) - J_{t+1}(z-1)))_+, \quad i = 1, \cdots, m.$$

Then by using relation (9), we have

$$H_t(z) = -\delta_t \sqrt{\sum_{i=1}^{m} \hat{p}_{it}^2 c_{it}^2 - \frac{\left(\sum_{i=1}^{m} \hat{p}_{it}^2 c_{it}\right)^2}{\sum_{i=1}^{m} \hat{p}_{it}^2}}.$$

Therefore, the robust counterpart of the dynamic programming formulation becomes

$$J_t(z) = J_{t+1}(z) + \sum_{i=1}^{m} \hat{p}_{it} c_{it} - \delta_t \sqrt{\sum_{i=1}^{m} \hat{p}_{it}^2 c_{it}^2 - \frac{\left(\sum_{i=1}^{m} \hat{p}_{it}^2 c_{it}\right)^2}{\sum_{i=1}^{m} \hat{p}_{it}^2}}, \tag{13}$$

where $1 \leq t \leq T$ and $0 \leq z \leq C$. Since the last term in (13) has an analytic solution, the computational complexity of the robust approach remains $O(mTC)$.

4. The Solution with Perfect Information

A useful concept in decision analysis is *perfect information.* Although this type of information rarely exists, it provides an upper bound on the value of real information since it pictures the "best case" scenario (see Clemen, R.T., T. Reilly (2003)). In our static problem setting, perfect information implies elimination of uncertainty about the total demand for each fare class. The subsequent model focuses on the perfect information from this "a priori" perspective. In Section 5, we solve the perfect information model approximately and compare our results with the results that we obtain after solving the other models of the previous sections.

Suppose that we decide on the allocation after knowing all the realized demands. Then, we obtain the following optimization model

$$v_4(C) := \mathbb{E}\left[\max\left\{\sum_{i=1}^{m} r_i \min\{x_i, \mathbf{D}_i\} \mid \sum_{i=1}^{m} x_i \le C, x_i \in \mathbb{Z}_+\right\}\right]. \tag{14}$$

It is obvious that $v_2(C) \le v_4(C)$. We may consider the positive difference $v_4(C) - v_2(C)$ as the expected cost of having *imperfect information.* We now introduce both the partial sum $\mathbf{S}_n := \sum_{j=n}^{m} \mathbf{D}_j$ with $\mathbf{D}_0 := 0$ and the stochastic process $\mathbf{N}(C) := \min\{0 \le n \le m \mid \mathbf{S}_n \le C\}$. Then by relation (2), the random optimal solution $(\mathbf{x}_i^*)_{i=1}^n$ for the random demands $\mathbf{D}_i$, $1 \le i \le m$ is given by

$$\mathbf{x}_i^* = \begin{cases} \mathbf{D}_i, & \text{if } i \ge \mathbf{N}(C) \\ C - \mathbf{S}_{\mathbf{N}(C)}, & \text{if } i = \mathbf{N}(C) - 1 \\ 0, & \text{if } i < \mathbf{N}(C) - 1. \end{cases}$$

The associated random optimal objective value equals

$$\mathbf{v}_1(C) = \sum_{i=\mathbf{N}(C)}^{m} r_i \mathbf{D}_i + \left(C - \mathbf{S}_{\mathbf{N}(C)}\right) r_{\mathbf{N}(C)-1}.$$

As in the deterministic case, for each realization this is a concave function in C. This shows that

$$v_4(C) = \mathbb{E}\left[\sum_{i=\mathbf{N}(C)}^{m} r_i \mathbf{D}_i + \left(C - \mathbf{S}_{\mathbf{N}(C)}\right) r_{\mathbf{N}(C)-1}\right]. \tag{15}$$

In general, it is difficult to give an analytical expression for this expectation, but we can approximate the above expectation by means of the Monte Carlo method (see Ross, S.M. (2002)).

5. Simulation Experiments

To support our theoretical study, we conduct simulation experiments and report our observations in this section. We compare, in the first subsection, the non-robust static model (3) with its robust counterpart (6). In the second subsection, a similar study is carried out to compare the non-robust dynamic model (12) with its counterpart (13). To see the differences between the static and the dynamic modeling approaches, we conduct additional simulation experiments in the final subsection. Using the same data, we also approximate the expectation in the perfect information model (14). We give then the comparison among static, dynamic and perfect information models. In all our simulation experiments we have used MATLAB 7.0 on a personal computer with 1.5 GHz Intel Celeron M processor and 256 MB of RAM.

In the following two subsections, we conduct simulation experiments to compare the robust and the non-robust models. Before discussing the simulation results, let us give the motivation of our setup. Consider an airline company, where the management tries to immunize the revenues against demand uncertainty. In many cases, the underlying distribution may not be obtained by statistical analysis (let alone its parameters) due to, for instance, insufficient or corrupted historical data. Therefore, the management may only provide an estimate ($\hat{p}$) and they may also guess the uncertainty set P in which the actual distribution (p) lies. Nevertheless, when reality reveals itself (a realization of p), the management can control the quality of their guess with parameter δ. In a sense, the choice of higher values of δ reflects the behavior of a more risk averse decision maker. Observe that the parameter set $\{\hat{p}, \delta\}$ plays a similar role to that of a prior distribution in Bayesian statistics. Using the motivation above, we give the main steps of our simulation experiments in Algorithm 1.

Algorithm 1 Main steps of The Simulation Experiments

1: Start with an estimated probability vector $\hat{p}$.

2: Using $\hat{p}$ solve the non-robust optimization model.

3: Using $\hat{p}$ and δ solve the robust optimization model.

4: Generate an actual distribution p from the uncertainty set P.

5: Generate *one* realization of the demand using p, and evaluate the revenues with robust and non-robust solutions obtained in steps 2 and 3, respectively.

6: Repeat steps 4 and 5, and collect the statistics (mean and standard deviations) of the total revenues for both non-robust and robust models.

7: Repeat steps 1 to 6 with different seeds.

As a final remark, notice that our setup is fundamentally different than the *conventional* approach. In the conventional approach, it is usually assumed that the general structure of the distribution may be obtained by statistical analysis using the historical data. However, the parameters of the distribution may be far away from the actual distribution parameters. Therefore, the error made by the estimation is mostly due to these faulty parameters. If we adopt this approach, then we should start with the actual distribution (p), and then simulate some historical data to find the estimated distribution ($\hat{p}$). However, as we discussed above, in our setup we start with the estimated distribution $\hat{p}$ and assume that when reality reveals itself, the actual distribution can be any distribution within a set around this estimated one.

5.1. Static Models: Non-robust vs Robust

We have implemented the algorithm given in Section 2.1. Recall that the same algorithm can also be applied to solve the robust version given in Section 2.2. As shown in relation (10) the convex subproblem has an analytic solution. Therefore, the only difference between the non-robust and robust implementations is the calculation of the $m \times C$ matrices given by (5) and (11), respectively.

We take 25 simulation runs with different seeds. As given in Algorithm 1, in each run we need to provide an estimated probability vector $\hat{p}_i \in \mathbb{R}^{K+1}$, $1 \leq i \leq m$. Then we use the algorithm discussed in Section 2.1 to find the optimal seat allocations of different fare classes for both the non-robust and the robust models. We next generate N realizations of the probability vectors $p_i \in \mathbb{R}^{K+1}$ uniformly from P_i, $1 \leq i \leq m$. The choice of uniform distribution conforms with our setup, since we assume, apart from the parameter δ, that the decision makers do not have information about the location of the actual distribution in the uncertainty set. Notice that to find these p_i vectors, one needs to generate uniform samples from the intersection of an ellipsoid and a hyperplane. This issue is discussed in Appendix A. After generating the probability vectors p_i, $1 \leq i \leq m$ as given in Algorithm 2, we then simulate the demand for each fare class using these probabilities. Next, the total revenues are evaluated according to non-robust and robust seat allocations. As our statistics, we store the mean and the standard deviation over N realized revenues in each run.

In actual applications, the estimated probability vectors $\hat{p}_i$ are provided by the decision makers. However, in our simulation experiments we also generate these estimated probability vectors using the truncated Poisson distribution with parameters $\lambda_i > 0$, $i = 1, \cdots, m$ and K. Consequently, the total demand for fare class i is concentrated on $\{0, \cdots, K\}$. In each run we select a set of λ_i values uniformly from the intervals $[\kappa_i, \mu_i]$, respectively. Then, we sort these values in descending order ($\lambda_1 > \lambda_2 > \cdots > \lambda_m$) to reflect the higher demand for relatively cheaper fare class seats. The parameters that we use are given in Table 1. An example of the estimated probability vectors obtained by using the truncated Poisson distribution is given in Figure 1.

Table 1 The parameters used in the simulation of static models.

Parameters	Values
$[N, K, C, m]$	[1000, 100, 100, 4]
(r_1, r_2, r_3, r_4)	(2, 3, 4, 6)
$(\kappa_1, \kappa_2, \kappa_3, \kappa_4)$	(40, 20 ,10 ,1)
$(\mu_1, \mu_2, \mu_3, \mu_4)$	(70, 40 ,30 ,10)

We report in Table 2 the simulation results over 25 runs. The figures in the table show the relative difference in percentages between the non-robust and the robust averages and standard deviations over N realized revenues for varying δ values. We elaborate these results in the accompanying Figure 2. The upper bar plot shows the relative difference in percentages between the non-robust

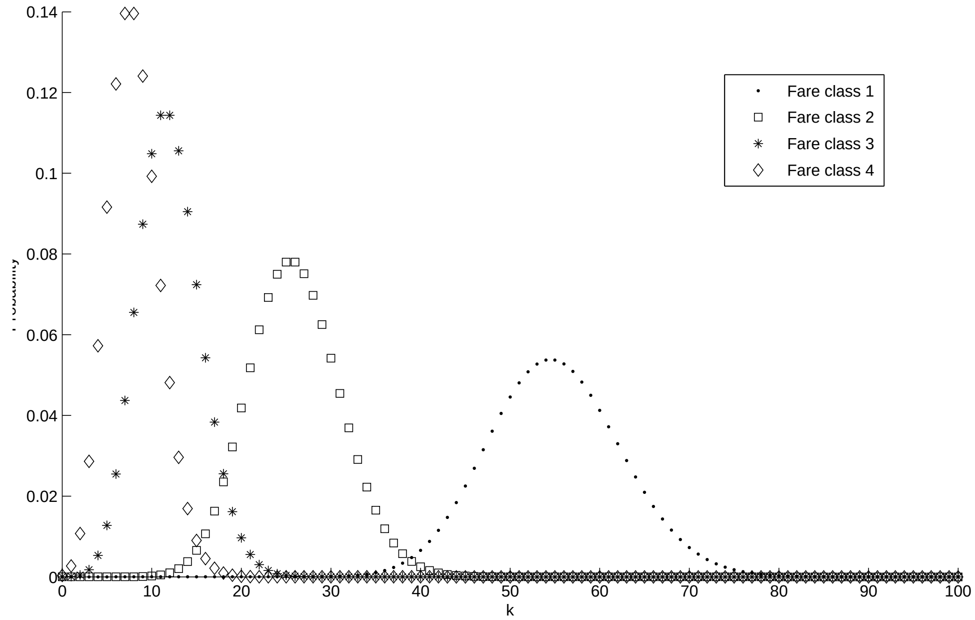

Figure 1 An example of the truncated probability distributions for different fare classes.

and the robust averages. This plot shows that even for the conservative choice of $\delta = 1$, the revenue obtained by the non-robust model in all runs is less than 0.35% above the robust model. On the other hand, the lower bar plot in Figure 2 shows the relative differences with respect to standard deviations. This plot shows that the deviation in the revenues obtained with the robust model is less than the deviation in the revenues obtained with the non-robust model. Therefore, we may assert that a stable solution is found with the robust model at the expense of a slight decrease in the revenue.

As shown in Figure 2, we also conduct sensitivity analysis with respect to δ values. The relative differences in means and standard deviations, in almost all cases, increase as δ increases. Notice that runs 8, 19 and 21 do not show any difference for any δ values because the optimal allocations in those runs are exactly the same. In fact, we observe the same phenomenon for smaller δ values (see, for example; runs 4, 12 and 15).

Since the convex subproblem has an analytic solution, the differences in computation times for robust and non-robust models is insignificant. Moreover, the simulation with the above parameters (for 25 runs) takes on average less than 3 minutes. Therefore, we do not report our computation

Table 2 The relative differences between robust and non-robust static models.

	Mean				Standard Deviation			
Runs	$\delta = 0.25$	$\delta = 0.50$	$\delta = 0.75$	$\delta = 1.0$	$\delta = 0.25$	$\delta = 0.50$	$\delta = 0.75$	$\delta = 1.0$
1	0.0000	0.0397	0.0383	0.2522	0.0000	5.1412	5.1311	14.5580
2	0.0383	0.0383	0.0394	0.1830	1.9532	1.9472	1.9496	3.6511
3	0.0530	0.0546	0.1557	0.1574	5.2615	5.2660	8.9313	8.9418
4	0.0000	0.0000	0.1627	0.1627	0.0000	0.0000	9.5396	9.5000
5	0.0000	0.0407	0.0393	0.0750	0.0000	8.1327	8.1177	11.7135
6	0.0000	0.0381	0.0381	0.1522	0.0000	4.8784	4.8686	14.3873
7	0.0000	0.1090	0.1115	0.2926	0.0000	8.4267	8.4336	15.8234
8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9	0.0000	0.0000	0.1199	0.1199	0.0000	0.0000	6.8531	6.8587
10	0.0141	0.0481	0.0496	0.2095	2.9007	7.8840	7.8881	16.8907
11	0.0776	0.1221	0.1221	0.1221	8.9355	16.5823	16.5681	16.5578
12	0.0000	0.0000	0.0560	0.0560	0.0000	0.0000	4.9875	5.0095
13	0.0000	0.0155	0.0667	0.1609	0.0000	4.4663	6.8890	16.2545
14	0.0000	0.0343	0.0605	0.2065	0.0000	5.3383	7.7911	15.3681
15	0.0000	0.0000	0.0628	0.2212	0.0000	0.0000	3.9593	13.8023
16	0.0000	0.1696	0.1717	0.1717	0.0000	11.5385	11.5142	11.5266
17	0.0000	0.0327	0.0727	0.2088	0.0000	4.5270	6.6342	13.9814
18	0.0000	0.1140	0.1111	0.1838	0.0000	13.2013	13.3008	16.4655
19	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
20	0.0641	0.0687	0.1787	0.1726	7.2190	7.3134	14.3019	14.2931
21	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
22	0.0000	0.0476	0.0490	0.2762	0.0000	5.1556	5.1746	15.5427
23	0.0000	0.0280	0.0266	0.2323	0.0000	6.3423	6.3436	14.5390
24	0.0000	0.0000	0.0845	0.0831	0.0000	0.0000	5.4288	5.3972
25	0.0325	0.0337	0.2210	0.2199	3.7505	3.7836	16.9918	16.9992

times separately. This remark is valid for all the subsequent results that we report.

5.2. Dynamic Models: Non-robust vs Robust

We have implemented a dynamic programming algorithm to solve (12). Since the convex subproblem of the robust model (13) has an analytic solution, only the calculation of the return at each stage is changed, and hence, the dynamic programming algorithm implemented for the non-robust model (12) is slightly modified to solve the robust version (13).

As in the previous subsection, we take 25 simulation runs with different seeds and in each run we provide the estimated probability vectors $\hat{p}_t \in \mathbb{R}^m$, $1 \leq t \leq T$. Then we compute the non-robust and the robust optimal policies by the corresponding dynamic programming algorithms. Using Algorithm 2 in Appendix A, we generate N realizations of the probability vectors $p_t \in \mathbb{R}^m$ uniformly from P_t, $1 \leq t \leq T$. Given a realization p_t, we simulate S times the arrival process, and then, using the non-robust and robust optimal policies, we compute the corresponding seat allocations. As our statistics, we store the mean and the standard deviation of the realized revenues.

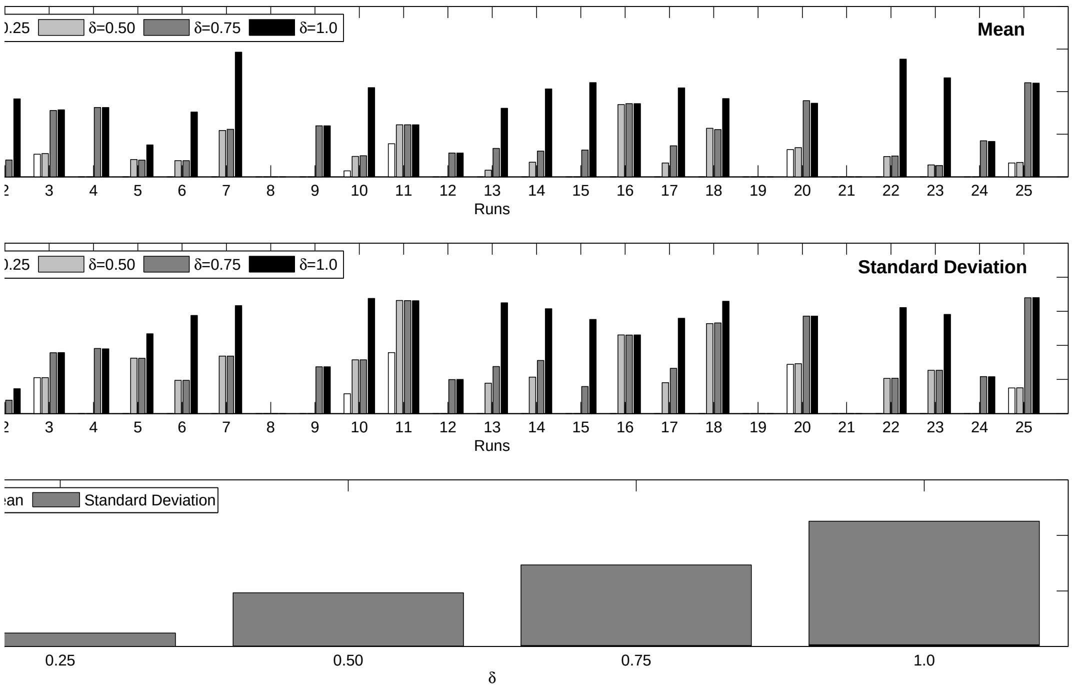

Figure 2 The relative differences between robust and non-robust static models with respect to means and standard deviations over N realizations for varying δ (bar plots). The average relative differences over 25 runs for each δ value (stacked bar plot).

In order to provide the estimated probability vector $\hat{p}_t$ of period t, we use a Dirichlet distribution with parameters $\gamma_i(t)$, $0 \leq i \leq m$. A Dirichlet distribution allows us to provide arrival probabilities at each period t for the fare classes. It is reasonable to predict that as the departure time T approaches, the requests for cheaper fare classes reduce, whereas the requests for the more expensive fare classes increase. To achieve this, we adjust the adopted Dirichlet distribution parameters monotonically. Figure 3 illustrates the change of these parameters over time. The actual values of the parameters that we use are given in Table 3.

Table 3 The parameters used in the simulation of dynamic models.

Parameters	**Values**
$[N, S, C, T, m]$	[100, 10, 100, 200, 4]
(r_1, r_2, r_3, r_4)	(2, 3, 4, 6)
$[\bar{v}_0, \bar{v}, v_0, v_1, v_2, v_3, v_4]^*$	[1, 2, 3, 5, 4 ,1, 0.5]

*See Appendix B for details.

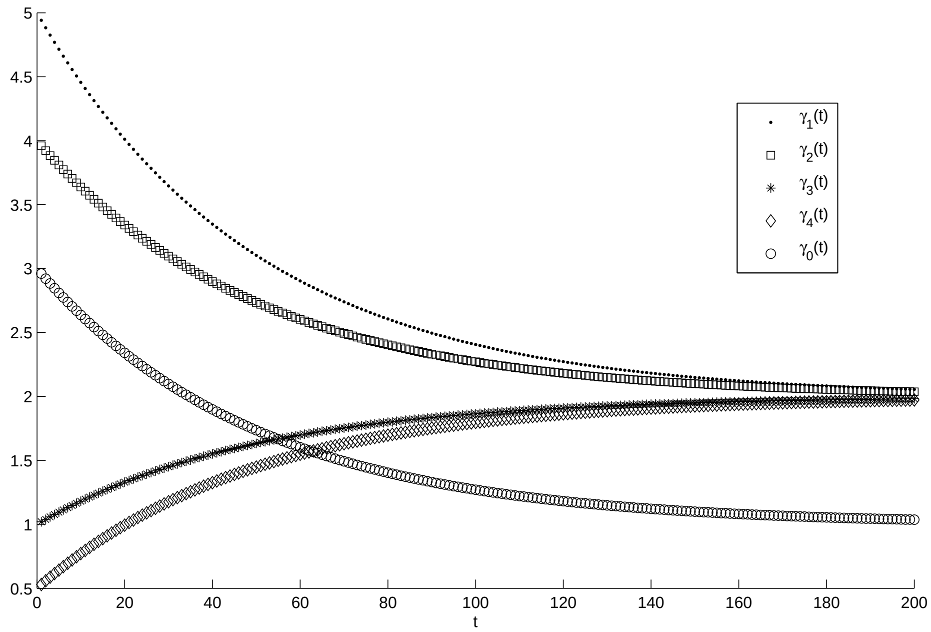

Figure 3 **The change of adopted Dirichlet distribution parameters over time.**

Similar to previous subsection, we report in Table 4 the simulation results over 25 runs. The figures in the table show the relative difference in percentages between the non-robust and the robust averages and standard deviations over N realized revenues for varying δ values. We work out the details of these results in Figure 4. The upper and lower bar plots represent, respectively, the means and the standard deviations over 25 runs as in Figure 2. In general, our results with the dynamic model is similar to the results obtained with the static model. In most cases, the non-robust models yield slightly better average revenues than the robust models (less than 1.5%). However, to our surprise, we find for runs 5, 10, 15 and 16 that when $\delta = 0.25$, the average revenues obtained with the robust model is barely larger than the non-robust model (hence the negative percentages in the plot). This happens especially when δ is small because for those values of δ the optimal policies of the two models yield almost the same allocations. The difference in two allocations may be just one seat allocated to a cheaper class ticket by the robust model. However, the policy suggested by the non-robust model fails to sell the one higher priced ticket for most of the realizations. The standard deviations plot in Figure 4 shows that the percentage deviation of the robust model is strictly less than the non-robust model in all runs even for small values of δ. As in the static case, we observe for most of the runs that both the averages and the standard deviations of the revenues increase as δ increases.

Shen, R., S. Zhang (2007) studied the properties of the solutions generated from a robust optimization model in the context of a multi-stage financial investment problem, and found that the solutions exhibit a substantially reduced variability in terms of the objective value, when the data in the model are subject to noises. Our findings in this paper, which is based on a very different decision model, have reconfirmed this remarkable property.

Table 4 **The relative differences between robust and non-robust dynamic models.**

	Mean				**Standard Deviation**			
Runs	$\delta=0.25$	$\delta=0.50$	$\delta=0.75$	$\delta=1.0$	$\delta=0.25$	$\delta=0.50$	$\delta=0.75$	$\delta=1.0$
1	0.1540	0.1714	0.6075	1.0218	1.7371	5.7958	11.7491	15.0325
2	0.0248	0.2558	0.7659	1.2637	0.6830	4.1842	12.1581	16.6348
3	0.0853	0.1812	0.5598	1.1210	1.5507	2.0598	10.3449	14.2507
4	0.0177	0.3752	0.6081	1.1031	3.5960	4.4298	10.3242	14.0830
5	-0.0242	0.1680	0.6499	1.2040	2.0314	7.5214	13.7690	15.3185
6	0.0804	0.0525	0.3508	0.7337	6.3103	2.7624	11.4959	13.6893
7	0.0082	0.2943	0.8156	1.3041	4.9740	8.5494	13.7976	17.2498
8	0.0939	0.2987	0.8010	1.3137	3.3910	5.6772	14.8163	18.4636
9	0.1298	0.1585	0.4576	0.8448	1.4964	5.8269	6.3663	11.4977
10	-0.0326	0.1780	0.6134	1.0601	1.4412	5.9982	13.7858	19.1628
11	0.0861	0.3199	0.5853	1.1140	3.2976	5.1494	9.7215	16.3041
12	0.0107	0.2051	0.5471	1.1384	4.2547	3.8342	9.1774	12.0738
13	0.0163	0.3427	0.7221	1.3797	6.2555	6.4410	8.5252	18.2730
14	0.1299	0.3547	0.7381	1.3060	1.1252	7.6007	12.2976	17.0356
15	-0.0317	0.0752	0.5860	1.1694	4.3253	6.3888	10.4889	20.4642
16	-0.0194	0.2222	0.3458	0.7691	6.8575	6.4745	9.6450	11.8082
17	0.0308	0.1652	0.3840	0.9279	2.8332	5.9426	6.9055	14.5609
18	0.1472	0.1542	0.4042	0.9098	3.6795	4.5889	12.3292	14.0818
19	0.1483	0.2668	0.6701	1.4077	1.8423	8.0660	11.3377	16.7693
20	0.0377	0.0286	0.4295	0.9092	1.2019	4.5848	10.6485	12.9858
21	0.0522	0.3706	0.5562	1.1848	1.0893	7.6126	10.7594	17.6085
22	0.1059	0.3025	0.5252	0.9221	4.3453	13.0542	10.2654	12.7708
23	0.1308	0.3089	0.6193	1.0713	1.2846	4.5705	7.5261	12.9965
24	0.0576	0.2482	0.5623	1.0452	4.2582	4.3260	7.7787	13.6414
25	0.1721	0.3703	0.6791	1.0448	4.3213	8.2487	10.2175	17.0594

5.3. Favorable Estimates

We now discuss the case when the estimated probability vector $\hat{p}$ coincides with the actual distribution p. Although unlikely, it may happen that the decision maker makes a very accurate guess for the estimated probability vector $\hat{p}$, and not knowing this *favorable estimate*, the decision maker

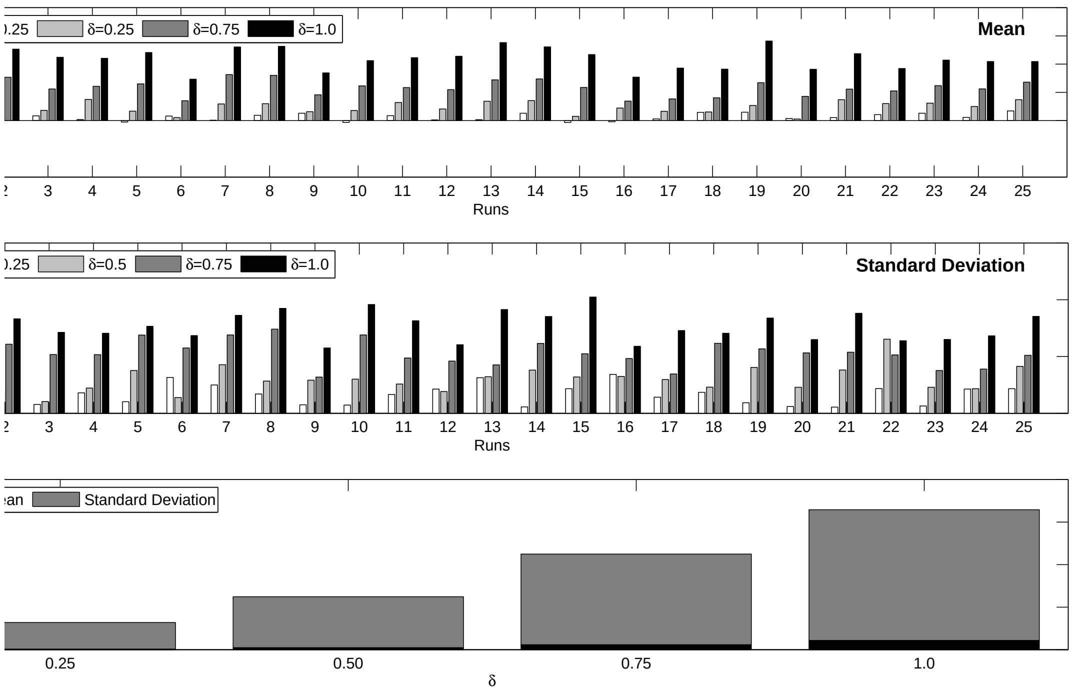

Figure 4 The relative differences between robust and non-robust dynamic models with respect to means and standard deviations over N realizations for varying δ (bar plots). The average relative differences over 25 runs for each δ value (stacked bar plot).

is still insecure about the estimation, and hence prefers to use a robust model with a given δ value. In the remaining part of this subsection, we conduct additional experiments for these favorable estimates cases regarding both static and dynamic (robust and non-robust) models. The following results can also be interpreted as relative losses between the robust and non-robust approaches in case the estimated probability vectors are, in fact, the actual distributions of the model.

The relative difference between the robust and the non-robust static models in terms of the averages and the standard deviations are listed in Table 5. These results are visualized in Figure 5. Using the same 25 seeds from Section 5.1, we first generate the truncated Poisson estimated probability distribution $\hat{p}$, $1 \leq i \leq m$ for the corresponding seed. However, unlike the generation from the uncertainty set P_i as in Section 5.1, we generate 1,000 realizations of fare class i demand from $\hat{p}_i$. This reflects the favorable estimate setting discussed in the previous paragraph, since the estimated probability distribution $\hat{p}_i$ coincides with the actual distribution vector p_i. Although there are slight differences between the results reported in Table 5 and the results obtained in Section 5.1 (see Table 2), on average these results are quite close and the differences are not significant. Hence, a similar assertion as in Section 5.1 follows; stable solutions are obtained with the robust approach at the expense of an admissible decrease in the revenues.

Table 5 The relative differences between robust and non-robust static models for favorable estimates.

	Mean				**Standard Deviation**			
Runs	$\delta = 0.25$	$\delta = 0.50$	$\delta = 0.75$	$\delta = 1.0$	$\delta = 0.25$	$\delta = 0.50$	$\delta = 0.75$	$\delta = 1.0$
1	0.0000	0.0185	0.0185	0.3078	0.0000	4.9813	4.9813	14.5815
2	0.0306	0.0306	0.0306	0.2021	2.0861	2.0861	2.0861	4.2066
3	0.0057	0.0057	0.0717	0.0717	5.3892	5.3892	8.8428	8.8428
4	0.0000	0.0000	0.1545	0.1545	0.0000	0.0000	8.4586	8.4586
5	0.0000	0.0179	0.0179	0.0704	0.0000	7.6564	7.6564	11.1085
6	0.0000	0.0214	0.0214	0.2063	0.0000	4.6223	4.6223	14.0976
7	0.0000	0.0520	0.0520	0.2235	0.0000	7.9708	7.9708	15.6421
8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9	0.0000	0.0000	0.1325	0.1325	0.0000	0.0000	7.6085	7.6085
10	0.0264	0.0626	0.0626	0.2261	2.2381	7.3258	7.3258	17.1703
11	0.1068	0.1550	0.1550	0.1550	9.1463	16.5994	16.5994	16.5994
12	0.0000	0.0000	0.0686	0.0686	0.0000	0.0000	4.8414	4.8414
13	0.0000	0.0655	0.1204	0.2733	0.0000	4.3112	6.8165	16.3515
14	0.0000	0.0289	0.0676	0.2330	0.0000	5.4740	7.8715	15.4315
15	0.0000	0.0000	0.0724	0.1755	0.0000	0.0000	4.3863	13.6209
16	0.0000	0.1367	0.1367	0.1367	0.0000	12.1906	12.1906	12.1906
17	0.0000	0.0122	0.0343	0.1902	0.0000	4.4968	6.9662	14.1078
18	0.0000	0.1198	0.1198	0.1975	0.0000	13.2627	13.2627	16.8737
19	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
20	0.0474	0.0474	0.1247	0.1247	7.4332	7.4332	14.2182	14.2182
21	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
22	0.0000	0.0599	0.0599	0.2723	0.0000	5.1861	5.1861	16.0540
23	0.0000	0.0308	0.0308	0.2217	0.0000	6.4055	6.4055	14.4105
24	0.0000	0.0000	0.0952	0.0952	0.0000	0.0000	5.9508	5.9508
25	0.0295	0.0295	0.2261	0.2261	3.7190	3.7190	17.2142	17.2142

A similar type of experiments as in Section 5.2 is conducted for the robust and non-robust dynamic models under the same assumption that the estimated probability of arrivals in each period coincide with the actual distribution of those arrivals. Again we do not generate the arrival process from a probability distribution in the uncertainty set, but from the estimated probability of arrivals (1,000 realizations of the arrival process are used). Under this scenario, the relative differences between the averages and the standard deviations of the robust and the non-robust dynamic models are listed in Table 6 and visualized in Figure 6. When we compare the results to those obtained in Section 5.2, our observations are similar: in most of the cases the non-robust

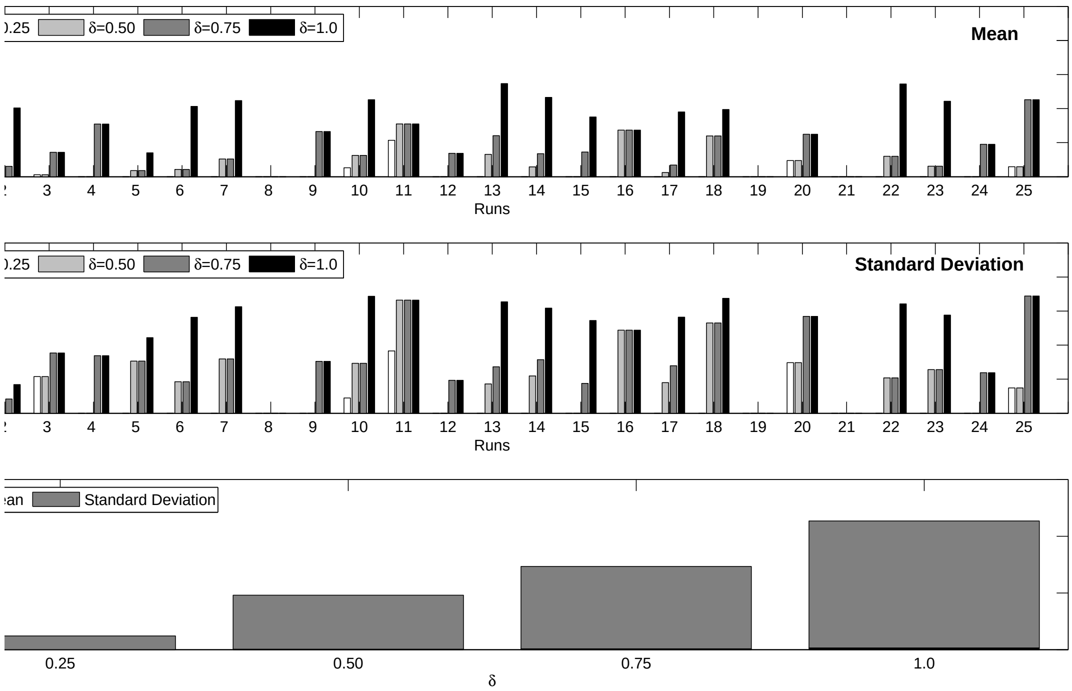

Figure 5 The relative differences between robust and non-robust static models in case of favorable estimates for varying δ (bar plots). The average relative differences over 25 runs for each δ value (stacked bar plot).

model yields slightly better revenues and the standard deviation of the robust model is less than the non-robust model. However, in three cases (runs 16, 18, 20) when $\delta = 0.25$, the non-robust models give smaller deviations. In those instances, the allocation with the dynamic model conforms with the realizations from the favorable estimate.

5.4. Cost of Imperfect Information

In this subsection we conduct simulation experiments to compare the static model (3), the dynamic model (12) and the perfect information model (14). The main motivation of these experiments is to check the effect of having additional information as one has more information in the dynamic model than the static model, and similarly, as the perfect information model includes more information than the dynamic model.

Again, we take 25 simulation runs with different seeds. In each simulation run, we first generate for $1 \leq t \leq T$ the arrival probability vector $p_t \in \mathbb{R}^m_+$ from a Dirichlet distribution with parameters γ_{it},

Table 6 The relative differences between robust and non-robust dynamic models for favorable estimates.

	Mean				Standard Deviation			
Runs	$\delta = 0.25$	$\delta = 0.50$	$\delta = 0.75$	$\delta = 1.0$	$\delta = 0.25$	$\delta = 0.50$	$\delta = 0.75$	$\delta = 1.0$
1	-0.0603	0.1059	0.3239	0.9246	0.0254	3.9757	13.1345	15.7227
2	0.1087	0.3122	0.6024	1.2370	4.5662	8.1873	11.3036	15.0279
3	0.0522	0.1027	0.4361	1.1137	3.8512	8.7411	12.8882	20.9504
4	-0.0827	0.3708	0.7027	1.1827	4.7887	6.8569	13.4421	15.5918
5	0.1438	0.4318	0.6834	1.2546	4.1562	6.8018	12.6940	18.7955
6	0.1682	0.2739	0.5248	0.9690	0.5447	4.9383	7.9241	13.7589
7	0.0994	0.1685	0.7886	1.2360	7.6623	11.8946	15.9216	20.8830
8	0.3048	0.7153	1.0457	1.7759	6.9537	14.1091	11.8425	20.4392
9	0.1535	0.1692	0.4564	0.9595	6.2172	6.5703	14.1544	17.6327
10	0.2823	0.4855	0.9184	1.6198	5.6858	12.4371	17.3933	17.6937
11	-0.0030	0.1762	0.5051	1.1443	4.8897	10.7907	15.2537	12.4340
12	0.1323	0.0707	0.5377	0.8604	1.5841	1.8624	8.0566	14.5105
13	0.1310	0.2657	0.6554	1.1191	4.4251	6.3187	10.8637	14.6125
14	0.3227	0.5186	0.7318	1.3539	6.8383	10.9298	11.8040	22.3445
15	-0.0212	0.2544	0.6595	1.1660	4.5884	11.4808	14.7982	19.4089
16	0.1981	0.2539	0.5061	0.7572	-2.8161	3.4697	5.9923	15.3736
17	0.0289	0.0466	0.2953	0.8493	5.5255	9.6808	13.2945	15.2775
18	-0.0030	-0.0206	0.2804	0.9513	-2.9443	0.7273	7.5608	10.9054
19	0.1189	0.3054	0.7446	1.4846	3.0497	6.0904	12.8275	16.0609
20	-0.0580	0.0613	0.4416	0.9389	-0.8882	2.0315	5.8633	9.2658
21	0.0824	0.3555	0.6850	1.1919	2.2579	1.7954	4.8239	12.1320
22	0.2862	0.1767	0.6863	0.9489	8.4060	9.0067	12.3898	18.8969
23	-0.2502	-0.1633	0.2992	0.8197	1.2970	6.5652	14.7672	16.7602
24	0.0347	0.2436	0.5659	1.0020	3.5711	9.4620	10.7577	13.8065
25	0.1203	0.2990	0.7020	1.1632	4.7246	8.7099	10.2037	13.2486

$0 \le i \le m$. As we discussed in Section 4, it is difficult to compute $v_4(C)$ and solve (14) to optimality. Therefore, we implemented a Monte Carlo algorithm, which generates N demand realizations according to p_t, $1 \le t \le T$, and then gives a point estimate of (15). Next, we compute the expected optimal revenue by the dynamic model (12). To make a fair comparison between the static and the other two models, we need to compute the demand probabilities $p_{ik} = \mathbb{P}\{\mathbf{D}_i = k\}$, $1 \le k \le T$, by using the arrival probabilities p_t, $1 \le t \le T$. Since $p_{it} = \mathbb{P}\{\xi_t = r_i\}$, $0 \le i \le m$, $1 \le t \le T$, we have

$$\mathbf{D}_i = \sum_{t=1}^{T} \mathbf{1}_{\{\xi_t = r_i\}}.$$

Since it is assumed that the random variables ξ_t, $1 \le t \le T$, are independent it follows that the Bernoulli random variables $1_{\{\xi_t = r_i\}}$, $1 \le t \le T$, are also independent. This shows for every $\alpha \in (0, 2\pi)$ that the discrete Fourier transform $\mathcal{P}(\alpha) = \mathbb{E}[\exp(\mathrm{i}\alpha \mathbf{D}_i)]$ has the form

$$\mathcal{P}(\alpha) = \mathbb{E}(\exp(\mathrm{i}\alpha(\sum_{t=1}^{T} \mathbf{1}_{\{\xi_t = r_i\}}))) = \Pi_{t=1}^{T} \mathbb{E}(\exp(\mathrm{i}\alpha \mathbf{1}_{\{\xi_t = r_i\}})).$$

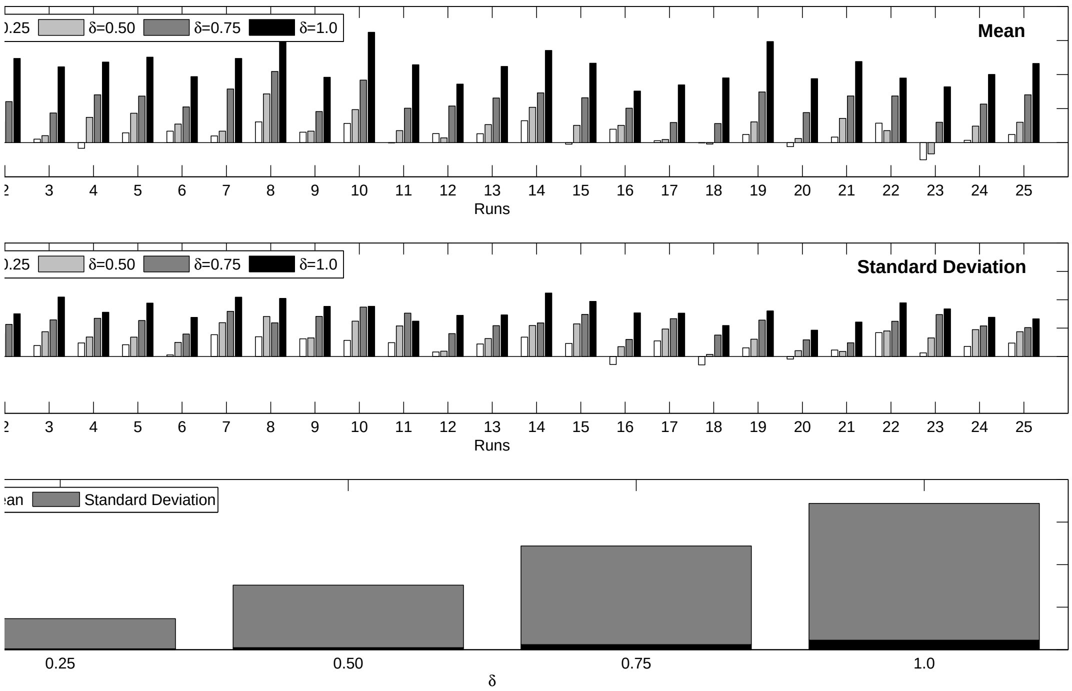

Figure 6 **The relative differences between robust and non-robust dynamic models in case of favorable estimates for varying δ (bar plots). The average relative differences over 25 runs for each δ value (stacked bar plot).**

Consequently,

$$\mathbb{E}\left[\exp(\mathrm{i}\alpha\mathbf{1}_{\{\xi_t=r_i\}})\right] = p_{it}\exp(\mathrm{i}\alpha) + (1-p_{it}) = 1 - p_{it}(1-\exp(\mathrm{i}\alpha))$$

and so, we obtain

$$\mathcal{P}(\alpha) = \Pi_{t=1}^{T}(1 - p_{it}(1-\exp(\mathrm{i}\alpha))).$$

It is well known that

$$p_{ik} = \frac{1}{T+1}\sum_{n=0}^{T}\mathcal{P}\left(\frac{2\pi n}{T+1}\right)\exp\left(\frac{-2\pi \mathrm{i}nk}{T+1}\right).$$

By using the FFT algorithm of the order $O(T\log T)$, one can easily recover the probabilities p_{ik} (see Golub, G.H., C.F. Van Loan (1996)). After recovering these probabilities, we compute the expected optimal revenue by the static model (3). As our statistics, we store the estimated total revenue of the perfect information model and the expected optimal revenues found by dynamic and static models, respectively. The parameters we use in our experiments are the same as in Table 3 except the parameter S, which is not required, and the parameter N, which is set to 1000.

Figure 7 shows our results as a stacked bar plot. The darker part of each bar in the plot shows the relative difference in percentages between the revenue obtained with the perfect information model (point estimate over N trials) and the revenue obtained with the dynamic model. Similarly, the lighter part of each bar denotes the relative difference in percentages between the perfect information and static models. Although the model with the perfect information yields higher revenues than both the dynamic and the static models, Figure 7 shows that the cost of imperfect information is rather insignificant when the dynamic model is considered. On the other hand, the cost of imperfect information increases as one prefers the static model over the dynamic one.

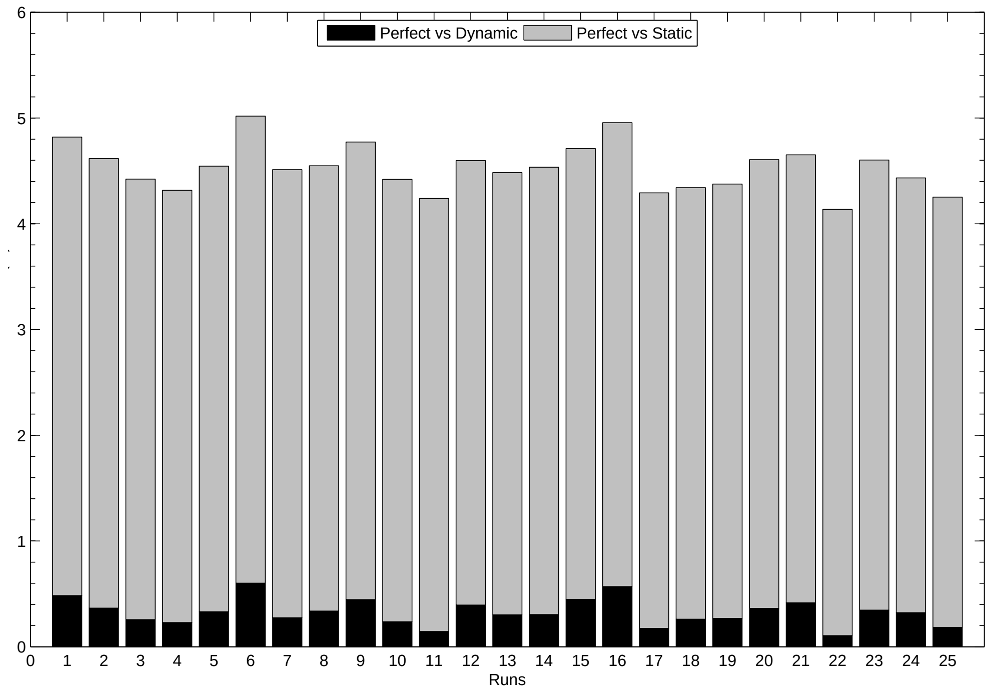

Figure 7 The relative differences between perfect information, dynamic and static models.

6. Conclusion

In this study we have shown by means of simulation that the use of robust versions of the classical static and dynamic single leg seat allocation problems in airline revenue management may be worthwhile due to the reduction in variability of the generated revenues. This reduction is much larger as the reduction in average revenue due to the conservative behavior of the considered robust models. In a subsequent paper we will consider extensions of the models in the network environment.

Appendix A: Uniform Sampling from The Uncertainty Set

Notice that in both static and dynamic model simulation runs, we need to generate sample vectors p_i, $1 \le i \le m$ and p_t, $1 \le t \le T$, from the intersection of an ellipsoid and a hyperplane of appropriate dimensions. In our subsequent discussion, we omit for ease of notation the subindices i and t.

To conduct our simulation experiments, we need to generate sample vectors p from the set

$$P = \{p \in \mathbb{R}^q_+ \mid p \in \hat{p} + \Delta, p^\intercal e = 1\},$$

where

$$\Delta = \left\{ x \in \mathbb{R}^q \mid \sum_{j=1}^{q} \left(\frac{x_j}{\hat{p}_j} \right)^2 \le \delta^2 \right\}.$$

Notice that $\hat{p}^\intercal e = 1$. Therefore, if we generate uniform samples from the set

$$S = \left\{ x \in \mathbb{R}^q \mid \sum_{j=1}^{q} \left(\frac{x_j}{\hat{p}_j} \right)^2 \le \delta^2,\ x^\intercal e = 0 \right\},$$

then we can set $p = \hat{p} + x$. Notice that S defines an ellipsoid on a $q-1$ dimensional subspace (see Figure 8). It is not straightforward to generate uniform samples from S. However, it is shown by Fang *et al.* that uniform samples can be easily generated from unit hyper-spheres (Fang, K.-T., S. Kotz, K.W. Ng 1990, Section 3.1.5). Therefore, we next apply two transformations so that we can transform S to a $q-1$ dimensional unit hypersphere.

Let $y = Ax$, where A is a $q \times q$ diagonal matrix with nonzero elements $(1/(\delta\hat{p}_1)), \cdots, 1/(\delta\hat{p}_q))$. Using this transformation, the set S becomes

$$S_y = \{y \in \mathbb{R}^q \mid y^\intercal y \le 1,\ y^\intercal \hat{p} = 0\}.$$

Since we want to focus only on the unit hypersphere, we further apply a transformation to reflect the vector $u := (\hat{p}/\|\hat{p}\|) - I_1$, where I_1 is the unit vector corresponding to the first column of the identity matrix I. This transformation is called Householder reflection (cf. Golub, G.H., C.F. Van Loan (1996)), and it is applied by using the orthonormal matrix

$$B = I - \frac{2}{u^\intercal u} u u^\intercal.$$

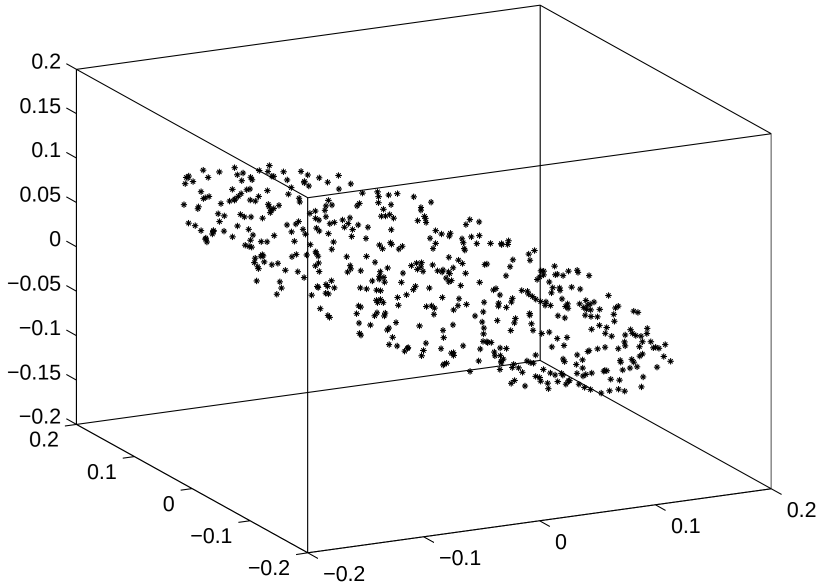

Figure 8 **A set of uniform samples from the ellipsoid centered at $\hat{p}^\intercal = (0.5, 0.2, 0.3)$ with $\delta = 1$.**

Using now $z = By$, the set S_y becomes

$$S_z = \{z \in \mathbb{R}^q \mid z^\intercal z \leq 1,\ z_1 = 0\}.$$

Notice that it is now enough to generate a realization of the vector $\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2, \cdots, \mathbf{Z}_q)$ uniformly from S_z. Then, using $B^{-1} = B^\intercal$ and the Jacobian transformation theorem, $\mathbf{X} = A^{-1}B^{-1}Z = A^{-1}B^\intercal \mathbf{Z}$ yields a uniformly distributed vector from S as desired.

To generate a realization of the vector $\mathbf{Z}$ from S_z, observe that we can equivalently generate a realization of the vector $\bar{\mathbf{Z}} = (\mathbf{Z}_2, \cdots, \mathbf{Z}_q)$ uniformly from the $q-1$ dimensional unit hypersphere

$$\bar{S}_z = \{z = (z_2, \cdots, z_q) \in \mathbb{R}^{q-1} \mid z^\intercal z \leq 1\}.$$

It is given (Fang, K.-T., S. Kotz, K.W. Ng 1990, pg. 75) that the random vector $\bar{\mathbf{Z}} = \mathbf{R}\mathbf{Q}$ is uniformly distributed on $\bar{S}_z$, where $\mathbf{Q}$ is a $q-1$ dimensional random vector distributed on the boundary of $\bar{S}_z$, $\mathbf{R}$ is a random variable with the distribution function

$$\mathbb{P}\{\mathbf{R} \leq r\} = r^{q-1},\ 0 \leq r \leq 1,$$

and the random variables $\mathbf{R}$ and $\mathbf{Q}$ are independent. Clearly by the inverse transformation method we obtain that $\mathbf{R} =^d \mathbf{U}^{(q-1)^{-1}}$ with $\mathbf{U}$ uniform distributed on $(0,1)$. To generate a realization of the random vector

$\mathbf{Q} = (\mathbf{Q}_1, \cdots, \mathbf{Q}_{q-1})$, we can generate for the components $\mathbf{Q}_i$, $1 \le i \le q-1$, independent standard normal variates and then normalize the resulting vector (cf. Fang, K.-T., S. Kotz, K.W. Ng (1990)). The following algorithm summarizes the steps to generate uniform samples from the set S. An illustrative set of samples generated by this algorithm is given in Figure 8.

Algorithm 2 Generating Uniform Samples from S

1: Generate standard normal variates $N_1, \cdots, N_{q-1}$ and a random number U.

2: Let $N = (N_1, N_2, \cdots, N_{q-1})$ and set

$$\bar{z} = \left(\frac{U^{(q-1)^{-1}} N_1}{\|N\|}, \ldots, \frac{U^{(q-1)^{-1}} N_{q-1}}{\|N\|} \right).$$

3: Set $z := \begin{pmatrix} 0 \\ \bar{z} \end{pmatrix}$ and return $x = A^{-1} B^{\intercal} z$.

Appendix B: Generating Arrival Probabilities for The Dynamic Models

In our simulation of the dynamic models, we generate the probability vectors $\hat{p}_t = (\hat{p}_{0t}, \hat{p}_{1t}, \ldots, \hat{p}_{mt})^{\intercal}$, $1 \le t \le T$ in the following way:

1. Generate some numbers $v_i, 0 \le i \le m$ and $\bar{v}_0, \bar{v}$ satisfying $0 < \bar{v}_0 < \bar{v} < v_0$, $0 < v_m < v_{m-1} < \ldots < v_1$ and $v_m < \bar{v} < v_1$.

2. Introduce the functions $\alpha_i : \mathbb{R}_+ \to \mathbb{R}$, $0 \le i \le m$ given by

$$\gamma_i(t) = v_i + (\bar{v} - v_i)(1 - \exp(-\frac{mt}{T})), \quad 1 \le i \le m$$

and

$$\gamma_0(t) = v_0 + (\bar{v}_0 - v_0)(1 - \exp(-\frac{mt}{T})).$$

3. Introduce the random vector $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_T) \in \mathbb{R}_+^{(m+1)\times T}$ consisting of the random vectors

$$\mathbf{X}_t = (\mathbf{X}_{0t}, \ldots, \mathbf{X}_{mt}), \ 1 \le t \le T$$

with the random variable $\mathbf{X}_{it}$, $0 \le i \le m$, $1 \le t \le T$ independent, and for each (i,t), the random variable $\mathbf{X}_{it}$ has a gamma distribution with scale parameter 1 and shape parameter $\gamma_i(t)$.

4. Introduce now for each (i,t)

$$\hat{p}_{it} = \frac{\mathbf{X}_{it}}{\sum_{j=0}^{m} \mathbf{X}_{jt}}.$$

It can be shown that the above procedure generates realizations $\hat{p}_t$ of a Dirichlet distributed random vector $\hat{p}_t$ with parameters $\gamma_0(t), \cdots, \gamma_m(t)$ (cf. Fang, K.-T., S. Kotz, K.W. Ng (1990)). This yields that

$$\mathbb{E}\left[\hat{p}_{it}\right] = \frac{\gamma_i(t)}{\sum_{j=0}^{m} \gamma_j(t)}.$$

Introducing now $i^* = \min\{1 \leq i \leq m \mid v_i > \bar{v}\}$ it follows by the definition of the function γ_i that the function γ_i is increasing for $i > i^*$ and decreasing for $i < i^*$. This modeling approach tries the capture the practical assumption that the arrival intensities are decreasing for the cheaper fare classes $i < i^*$ in the total remaining time before departure of the plane (but always above the arrival intensities of the more expensive fare classes $i \geq i^*$), while for the more expensive fare classes $i \geq i^*$ are increasing in the remaining time before departure. Observe for t large enough and $1 \leq i \leq m$ that

$$\mathbb{E}\left[\hat{p}_{it}\right] \approx \frac{\bar{v}_i}{\sum_{j=0}^{m} \bar{v}_j}$$

and $t \mapsto \mathbb{E}\left[\hat{p}_{it}\right]$ is increasing in t for $i > i^*$ and decreasing for $i \leq i^*$.

Acknowledgments

We thank the anonymous referees for their helpful suggestions and comments.

References

Belobaba, P.P. 1989. Applications of a probabilistic decision model to airline seat inventory control. *Operations Research* **37** 183–197.

Ben-Tal, A., A. Nemirovski. 1998. Robust convex optimization. *Mathematics of Operations Research* **23** 769–805.

Ben-Tal, A., L. El Ghaoui, A. Nemirovski, ed. 2006. *Robust Optimization*, *Mathematical Programming, Ser. B*, vol. 107.

Bertsekas, D.P. 1999. *Nonlinear Programming (2nd Edition)*. Athena Scientific, Belmont.

Bertsimas, D., M. Sim. 2003. Robust discrete optimization and network flows. *Mathematical Programming* **98** 49–71.

Brumelle, S.L, J.I. McGill. 1993. Airline seat allocation with multiple nested fare classes. *Operations Research* **41** 127–137.

Clemen, R.T., T. Reilly. 2003. *Making Hard Decisions with Decision Tools, 2nd Edition*. Duxbury.

De Boer, S.V., R. Freling, N. Piersma. 2002. Mathematical programming for network revenue management revisited. *European Journal of Operational Research* **137**(1) 72–92.

Fang, K.-T., S. Kotz, K.W. Ng. 1990. *Symmetric Multivariate and Related Distributions*. Chapman and Hall, New York.

Golub, G.H., C.F. Van Loan. 1996. *Matrix Computations, 3rd Edition*. The Johns Hopkins University Press, Baltimore.

Lautenbacher, C.J., S. Stidham Jr. 1999. The underlying Markov decision process in the single-leg airline yield-management problem. *Transportation Science* **33**(2) 136–146.

Lee, T.C., M. Hersch. 1993. A model for dynamic airline seat inventory control with multiple seat bookings. *Transportation Science* **27** 252–265.

Littlewood, K. 1972. Forecasting and control of passengers. *Proceedings of the 12th Annual AGIFORS Symposium* 95–117.

McGill, J.I., G.J. Van Ryzin. 1999. Revenue management: research overview and prospects. *Transportation Science* **33**(2) 233–256.

Richter, H. 1982. The differential revenue method to determine optimal seat allotments by fare type. *Proceedings of the 22nd Annual AGIFORS Symposium* 339–362.

Robinson, L.W. 1995. Optimal and approximate control policies for airline booking with sequential non-monotonic fare classes. *Operations Research* **43** 252–263.

Ross, S.M. 2002. *Simulation (3rd Edition)*. Academic Press, New York.

Shen, R., S. Zhang. 2007. Robust portfolio selection based on a multi-stage scenario tree. *European Journal of Operational Research* **doi:10.1016j/ejor.2007.01.059**.

Talluri, K.T., G.J. Van Ryzin. 2004. *The Theory, Practice of Revenue Management*. Kluwer Academic Publishers, Dordrecht.

Walczak, D., S. Brumelle. 2003. Dynamic airline revenue management with multiple semi-markov demand. *Operations Research* **51** 137–148.

Wollmer, R.D. 1986. A hub-spoke seat management model. Tech. rep., Mc Donell Douglas Co., Long Beach, CA.

Wollmer, R.D. 1992. An airline seat management problem for a single leg when lower fare classes book first. *Operations Research* **40** 26–37.