November, 2012

Revisiting the Newsboy Problem-Optimization with a Little Help from the Airline Industry

Tamas Lengyel, *Occidental College*

Revisiting the Newsboy Problem-Optimization with a Little Help from the Airline Industry

Tamás Lengyel

Department of Mathematics, Occidental College, 1600 Campus Road, Los Angeles, CA90041, USA

lengyel@oxy.edu

Abstract-In a typical inventory planning problem with a life cycle of only one planning period, we incur the cost of production per unit produced, profit per unit sold, loss per unit not sold, and lost revenue per unit ordered but not matched due to the lack of availability. The goal is to find the inventory level that maximizes the expected net profit. Textbooks often use the newsboy problem to illustrate the inventory management paradigm. The derivation of the formulas for the optimal level is usually done on an ad hoc basis, by dull and rote mathematical manipulations, for each modification of the simple basic model. The only purpose of this note is to give a simple transparent proof of the fact that quite surprisingly the lost revenue can be combined with the profit by reducing the general problem to a well known simplified case with no lost revenue. The reduction uses an airline analogy and thus, with some tweaking, it places the proof into a classical revenue management paradigm. We also provide an alternative derivation of the optimal solution for the discrete case which integrates the problem into a much broader class of optimization problems.

Keywords- Newsboy Problem; Inventory and Revenue Management; Optimization

I. INTRODUCTION

The newsboy (or newspaper boy or news-vendor, etc.) problem is a classical example in inventory management, and it can be traced back to Edgeworth's work in 1888. Although in this note we have no interest in surveying the literature, we mention that an extensive historical overview is given in [1], [2] and [3]. A popular textbook with a reasonable introductory coverage of inventory theory can be found in [4].

The newsboy has to make a decision on how many newspapers to carry. If he stocks up too many copies then he will be left with unsold publications that have no value at the end of the day. If he carries too few copies then some customers will be unsatisfied. The problem's main goal is to optimize the expected net profit by finding and setting the appropriate stock level. The use of expected value is generally justified by the law of large numbers, cf. [2].

A similar situation arises when managers make decisions about inventory levels of seasonal goods, such as Christmas cards that should satisfy the demand in December. Any cards left over in January have only a small residual value. This single-period model is also often used in the case of perishable goods and the fashion and apparel industries. Moreover, there is a downward trend in the life cycles of products in service industries and high-tech retail, and it leads to the growing importance of this model and its extensions as mentioned in [3].

Another more involved application is airline booking. Having empty seats corresponds to having too many newspapers. On the other hand, if there are passengers at the gate who can't get on the plane then that corresponds to too few newspapers. It is customary to offer a bumping reward to the latter quite disappointed passengers to compensate for the inconvenience and as a gesture of goodwill. Sometimes the reward is in the form of a voucher to be used with the same airline which may generate future business and thus, reduce the actual loss.

One obvious difference from the newsboy problem is that the stock level is set by the actual number s of available seats and therefore, the demand is censored by the number n of reservations taken. Of course, by setting n we can determine the optimal s, and if it is not equal to the given s then we can find the right n. Another potential difference is that since passengers pay different fares on the same flight, the profit per passenger may also vary. In the case of nonrefundable fares, no-shows forfeit the fare and thus, always contribute to the net profit (besides leaving room for more passengers). We will ignore these last two possibilities, though interested readers might consult [5], one of the authoritative monographs on revenue management, or [4, Section 18.8] on overbooking related issues that are typically more complex than the one we need here. We note that revenue management systems are getting increasingly popular in service industries, e.g., hotels, car rental companies, tour operators, etc.

The only purpose of this note is to give a simple and transparent proof of the fact that when optimizing the expected net profit, we can combine two seemingly antipodal factors, the lost revenue and the profit. This is accomplished by changing the model with the lost revenue incorporated into the profit–but only on a theoretical level without changing other factors, e.g., the demand distribution which might be affected if we simply increased the profit. In this way the reduction is achieved by a reasonable change in the model rather than the customary mathematical derivation which does not seem to shed light on the reason why the simplified interpretation is possible. After the combination (cf. Theorem III.1), we can reduce the general problem to a well known simplified case with no lost revenue. We also provide an alternative proof for the discrete case that shows connections to more general optimization problems, cf. Problem in Section III.

II. THE ACTUAL PROBLEM

Let X denote the random variable with distribution D and probability density function $f(x)$ of the demand in certain units, e.g., dollars.

We address the general case which involves different types of cost, profit, and losses. In Section III, we present a uniform approach in which some of these quantities can be effectively combined. Let c, g, l, and r stand for the cost of production per unit produced, price (gain) per unit sold (with $g \geq c$), loss per unit not sold, and lost revenue (or bumping reward in the airline context) per unit ordered but not matched due to the lack of availability. Everything is measured in dollars. Let $W_s(c,g,l,r,D)$ and $EW_s(c,g,l,r,D)$ denote the random variable corresponding to the total profit and its expected value, respectively, for the seller if s units are stocked. This quantity ultimately depends on the demand X which is independent of s and the number of sales V_s with s units stocked.

In the case of overstocking the profit is $gX-l(s-X)-cs= (g+l)X-(l+c)s$, if $X < s$, and in the case of under stocking it is $gs-r(X-s)-cs = (g+r-c)s-rX$, if $X > s$. Note that both forms work if the stock level s is properly set at X. We also add that if $c = 0$ then $l > 0$ represents the loss due to unsold items. If $c > 0$ and $l < 0$ then $-l$ may correspond to the per unit salvage value. On the other hand, if $c > 0$ and $l > 0$ then l may represent extra cost due to restocking and storage. In general, $c+l \geq 0$ is the actual per unit cost combined with loss and salvage due to unsold units.

For example, we want to find the maximum expected profit for the demand $X \sim Binomial[n = 10, p = .50]$, with $g = \$3$, $l = \$1$, and $c = r = \$0$. After graphing $EW_s(0, g, l, 0, Binomial[n = 10, p = 0.50])$, $s = 1, 2, \ldots, 10$, we obtain that the best choice for s is 6. Note, however, that we can find the answer without any graphing by finding the $g/(g+l) = 3/(3+1) = 0.75$ quantile value of the distribution function of the demand as stated in Theorem III.2.

III. THE REDUCTION

Our goal is to present a method that reduces the general problem to its most well known base case of the newsboy problem with $c = r = 0$, see e.g., [6, Example 4b, pp145-146]. Of course, the cost c can be easily introduced into this case. However, it is somewhat surprising that r can be absorbed by g. This fact is quite counterintuitive since we don't expect the per unit loss r to be combined with the per unit gain g. As we will see, we can reduce the discussion to calculations with $W_s(0, g+r, l, 0, D)$. Although, the formulas are well known in the general case, and usually derived by dull mathematical manipulations, we have not found an explanation or suggestion in the literature for such a simple reduction.

Besides the cost of production, we have one source of gain and two sorts of losses. We might encounter loss due to leftover units and loss due to losing business (or compensating for inconvenience, e.g., offering bumping rewards in the airline business). In both cases, the loss depends on how the actual level of demand compares to the level of stocking. Amazingly, as we mentioned, the second kind of loss can be combined with the gain. To prove this we use the airline analogy. Since the argument involves an extra charge paid by the hopeful passengers, it might turn out to be quite appealing to the airlines but a rather dangerous mental exercise from the customer's point of view–not to mention that once implemented passengers might be sensitive to higher ticket prices.

In fact, the airline industry has always been the most creative in embracing new ideas for increasing revenue. As the Los Angeles Times reported in its Daily Travel & Deal Blog in September of 2009 (cf. http://travel.latimes.com/daily-deal-blog/index.php/southwest-airlines-a-5256), Southwest Airlines added "an optional charge for 'EarlyBird Check-In,' the right to board the plane immediately following Southwest's Business Select and Rapid Rewards A-List customers. Fliers can pay an extra $10 for the peace of mind that they'll get to board as soon as possible and grab an open seat just that much sooner. Also included in the service is automatic check-in within 36 hours of your flight's departure". Southwest also added extra fees for unaccompanied minors, dogs and cats, and doubled its fees for a third or overweight bag in 2009. Other companies use checked-bag charges (cf. http://www.airfarewatchdog.com/blog/3801089/airline-baggage-fees-chart/).

The proof suggests an extra fee for each potential passenger. We have no doubt that, by making an appealing mathematical explanation go terribly wrong, the airlines will be happy to explore this surcharge option as well. On the other hand, we are not sure that the obvious educational benefits gained from simplifying the optimization problem will compensate for any fee and think that the airline industry should not listen...

Theorem III.1. We have that

$$W_s(c, g, l, r, D) = W_s(0, g + r, l, 0, D) - cs - rX. \qquad (1)$$

***Proof*:** Clearly,

$$W_s(c, g, l, r, D) = W_s(0, g, l, r, D) - cs,$$

and thus, without loss of generality and after subtracting the cost of production cs, we suppose that $c = 0$. Now we borrow the terminology from the airline context, and assume that potential customers are required to pay an entrance fee upon arrival at the airport of their departure. This fee of $\$r$ is paid by all passengers who present themselves at check-in, irrespective of whether they will be accommodated or not. Note that the concept of the entrance fee is different from that of paying the fare in advance since no-shows do not pay this fee.

In comparing the two sides of (1), we can ignore the "profit" due to the loss l per unit unsold since it is common in both "models," and hence the difference $W_s(0, g+r, l, 0, D) - W_s(0, g, l, r, D)$ is not affected.

We consider the quantity $W_s(0, g, l, r, D) + rX$. In fact, we observe that the total profit comes from selling X units and generating $g+r$ profit per unit sold if the demand can be completely satisfied, i.e., $X \leq s$. However, if this is not the

case, then the bumping reward wipes out the entrance fee for every order beyond s, and the only gain that remains is $(g+r)s$ from the first s orders.

In the case of $W_s(0, g+r, l, 0, D)$, a $g+r$ profit is generated for every unit sold if $X \le s$ as in the previous case. On the other hand, now there is no bumping reward; thus, a per unit profit of $g+r$ is generated for each of the first s satisfied customers if $X > s$. We get that the total profit $W_s(0, g + r, l, 0, D)$ is equal to that of the first model.

For the sake of completeness, we derive a result about finding the best value of s which is a generalization (cf. [7, pp113-114] or [4, S18.7, pp875-877]) of a well-known fact in the case of $c = r = 0$ (cf. [6, Example 4b, pp145-146]).

Theorem III.2. Let $F^{-1}(x)$ denote the quantile function, i.e., the inverse of the cumulative distribution function $F(x)$ of X in the sense that $F^{-1}(x)$ is equal to the smallest y such that $F(y) = x$. Then we have that

$$s = F^{-1}\left(\frac{g+r-c}{g+r+l}\right) \tag{2}$$

for the number s of units to be stocked to optimize the expected profit.

Proof: We use Theorem III.1 and note that to maximize the expected profit $EW_s(c,g,l,r,D)$ with respect to s we can ignore the term $-rX$ in (1) since it does not depend on s.

First we deal with the case in which X is a discrete random variable. We derive that the expected value of $W_s(c,g+r,l,0,D)$ is

$$EW_s(c,g+r,l,0,D) = (g + r)\sum_{i\le s} iP(X = i)$$

$$+(g + r)s\sum_{i>s} P(X = i) - l\sum_{i\le s}(s-i)P(X = i) - cs,$$

and after adjusting the third term on the right hand side by $-l\sum_{i>s}(s-s)P(X = i) = 0$, we get that

$$EW_s(c,g+r,l,0,D) = (g+r)\, EV_s - l(s-EV_s) - cs \tag{3}$$

$$= (g+r+l)\, EV_s - (l+c)s,$$

with V_s denoting the number of sales.

We note that (3) can be easily derived without any calculation: aside of the production cost, the gain is due to the number of sales and the loss comes from overstocking. The latter quantity is l times the size s reduced by the expected size of the sales since sales do not generate any loss.

We also get that the difference in expected total profit by preparing for one more customer is $\Delta_{s+1} = EW_{s+1}(0, g+r, l, 0, D) - EW_s(0, g+r, l, 0, D) = (g+r+l)E(V_{s+1} - V_s) - (l+c) = (g+r+l)\sum_{i\ge s+1} P(X = i) - (l+c)$ since $V_{s+1} - V_s$ is the indicator variable of the event that $X=i\ge s+1$ if X is a discrete random variable. Thus the expected profit reaches its largest value if s is largest so that

$$\Delta_{s+1} = (g + r + l)\sum_{i\ge s+1} P(X = i) - (l + c) > 0$$

since clearly, Δ_s is a decreasing function in s. The solution is one more than the largest s for which

$$\sum_{i\le s} P(X = i) < \frac{g+r-c}{g+r+l},$$

or equivalently, the smallest s so that

$$P(X \le s) \ge \frac{g+r-c}{g+r+l}.$$

If X is a continuous random variable then first we observe that

$$\frac{\partial}{\partial s} EV_s = \frac{\partial}{\partial s}\left(\int_0^s xf(x)dx + s\int_s^\infty f(x)dx\right) = P(X \ge s).$$

Similar to our use of (3) in deriving Δ_{s+1}, this yields that

$$\frac{\partial}{\partial s} EW_s(c, g+r, l, 0, D) = (g + r + l)P(X \ge s) - (l + c)$$

which is positive as long as

$$P(X \le s) \ge \frac{g+r-c}{g+r+l}$$

as in the discrete case.

For a given demand distribution, the value of s can be easily found by, say, using some software package, e.g., S-PLUS or R. From a statistical point of view, if F is continuous and unknown then we can take the order statistics of a sample of size n from the demand distribution and construct confidence intervals of the form $[X_{i:n}, X_{j:n}]$ for any required quantile. Here i and j depend on the given confidence level but not on the actual distribution; thus, it provides a distribution-free estimation method for s. If F is discrete then this interval will work at the given or higher confidence level, cf. [8].

The solution given in (2) tells us that we should set the stock level to satisfy $(g + r - c)/(g + r + l)$ fraction of the demand. To interpret this we can benefit from switching from maximizing expected profit to minimizing expected cost. As a rule of thumb, we can view the numerator $g - c + r$ as the unit cost of under ordering (or "underage"), i.e., decrease in net profit due to failing to order a unit that could have been sold, including loss of customer goodwill, and the denominator as the sum of this cost and the unit cost $c+l$ of over ordering (or "overage"), i.e., decrease in net profit due to ordering a unit that could not be sold, cf. [4]. This can be justified by noting that the discrete case can be easily treated as a special case of the following

Problem (Problem 2 in [9]) Let $x_1 < x_2 < \ldots < x_n$ be n real numbers. Given the positive weights $w_1, w_2,\ldots,w_n$ find an a such that the minimum of

$$D(a) = D(a; w_1, w_2,\ldots,w_n) = \sum_{i=1}^{n} w_i |x_i - a|$$

is achieved.

We rephrase the solution from [9]. The function $D(a)$ is non-negative, continuous and piecewise linear, so its minimum is attained at one of the points where the linear segments are joined, i.e., at some x_m. This problem can easily be reduced to finding the minimum m such that

$$w'_m = \sum_{i=1}^{m} w_i \Big/ \sum_{i=1}^{n} w_i \geq 0.5 \qquad (4)$$

(The optimum value $a = x_m$ is sometimes referred to as the weighted median of the values x_i with weights w_i.) To see this we need only to check the changes in $D(a)$ as a moves from the left of x_1 to the right of x_n. A change from a to $a+h$, $h > 0$, within the interval $[x_m, x_{m+1})$ yields the change $D(a+h)-D(a) = h\left(\sum_{i=1}^{m} w_i - \sum_{i=m+1}^{n} w_i\right)$; thus, we should increase a until $D(a+h)-D(a)$ ceases to be negative, and thus, (4) follows.

To apply this to our situation in which we have a discrete demand distribution with a finite support set of size n, we set $x_i = i - 1$, $1 \leq i \leq n$, $w_i = (c + l)P(X = x_i)$ for $1 \leq i \leq m = s+1$ and $w_i = (g-c+r)P(X = x_i)$ for $s+1 < i \leq n$. The criterion (4) turns into $w'_m = (c+l)F(s)/((c+l)F(s)+(g-c+r)(1-F(s))) \geq 0.5$, i.e., $F(s) \geq (g-c+r)/(g+r+l)$. Clearly, the problem and its solution can be generalized to an infinite support set $x_1 < x_2 < \ldots$ as long as $\sum_{i=1}^{\infty} w_i$ is finite.

IV. DISCUSSION AND CONCLUSIONS

Textbooks often use the newsboy problem to illustrate the inventory management paradigm. They derive the optimum under different settings by the application of the same approach: calculate the benefit of slightly changing the inventory level. This standard approach becomes a repetitive task requiring only pedagogically counterproductive rote calculations. In this note we found that determining the optimum tradeoff between over and under stocking in the various inventory settings should not be a cumbersome mathematical task. In fact, introducing the cost of lost revenue into the usual basic model does not lead to any complications if one uses the right reduction. We have not found an explanation or suggestion in the literature for such a simple reduction. In addition, the underlying optimization problem can be rephrased in terms of more general problems which make the process and result of solution more transparent.

It would be interesting to explore other more complex models whether similar reduction could be applied for finding optimum solutions.

ACKNOWLEDGMENT

The author wishes to thank Gregory P. Tollisen for helpful suggestions and comments.

REFERENCES

[1] Evan L. Porteus, Stochastic inventory theory. In Stochastic models, volume 2 of Handbooks Oper. Res. Management Sci., pages 605–652. North-Holland, Amsterdam, 1990.

[2] H. L. Lee and S. Nahmias, Single-product, single-location models. In Logistics of production and inventory, volume 4 of Handbooks Oper. Res. Management Sci., pages 3–55. North-Holland, Amsterdam, 1993.

[3] M. Khouja, The single-period (news-vendor) problem literature: review and suggestions for future research. Omega, 27(5):537–553, 1999.

[4] Frederick S. Hillier and Gerald J. Lieberman, Introduction to operations research, McGraw-Hill Inc., New York, NY, ninth edition, 2010.

[5] Kalyan T. Talluri and Garrett J. van Ryzin, The theory and practice of revenue management, International Series in Operations Research & Management Science, 68. Kluwer Academic Publishers, Boston, MA, 2004.

[6] Sheldon Ross, A first course in probability, Pearson Prentice Hall, New York, eighth edition, 2010.

[7] E. Kalvelagen and H. C. Tijms, Exploring operations research and statistics in the micro lab, Prentice-Hall Inc., Upper Saddle River, NJ, 1990.

[8] H. N. Nagaraja, Order statistics from discrete distributions. Statistics, 23(3):189–216, 1992. With discussion and a rejoinder by the author.

[9] Gregory P. Tollisen and Tamás Lengyel. On minimizing distance by the road less traveled, Elem. Math., 58(3):89–107, 2003.

Tamás Lengyel was born in Budapest, Hungary. He graduated as mathematics major with emphasis on Operations Research from Eötvös University, Budapest, Hungary in 1974 and received his Ph.D. in mathematics (probability theory and statistics) with a thesis on combinatorial and probability theoretic problems in cluster analysis from the same university in 1979.

He worked at the Computer and Automation Research Institute of the Hungarian Academy of Sciences, Budapest, Hungary as a Research Associate then Research Fellow and finally, Head of Department. He was a Visiting Professor of Mathematics and Computer Science at Beloit College, WI, USA. Currently, he is Professor of Mathematics at Occidental College, CA, USA.

He has published more than 60 papers, including 10 articles on the various properties of the Stirling numbers of the second kind. The Lengyel constant is named after him by Profs. Flajolet and Salvy. His current research interest is in the combinatorial, algorithmic, and number theoretical aspects of probability, statistics, applied mathematics, and game theory.

Prof. Lengyel is a member of the American Mathematical Society and the János Bolyai Mathematical Society. He received the Gyula Farkas Memorial Prize of the Hungarian Academy of Sciences and has been a referee and reviewer for academic programs, the NSA, many mathematical publications and editor of the quarterly journal Advances and Applications in Statistics.