Use of Uncertain Additional Information in Newsvendor Models

1st author: Sergey Tarima
Institute for Health and Equity
Medical College of Wisconsin
Wauwatosa, U.S.A.
starima@mcw.edu

2nd author: Zhanna Zenkova
Institute of Applied Mathematics and Computer Science
Tomsk State University
Tomsk, Russia
zhanna.zenkova@mail.tsu.ru

2020 5th International Conference on Logistics Operations Management (GOL) | 978-1-7281-6425-0/20/$31.00 ©2020 IEEE | DOI: 10.1109/GOL49479.2020.9314751

***Abstract*—The newsvendor problem is a popular inventory management problem in supply chain management and logistics. Solutions to the newsvendor problem determine optimal inventory levels. This model is typically fully determined by a purchase and sale prices and a distribution of random market demand. From a statistical point of view, this problem is often considered as a quantile estimation of a critical fractile which maximizes anticipated profit. The distribution of demand is a random variable and is often estimated on historic data. In an ideal situation, when the probability distribution of demand is known, one can determine the quantile of a critical fractile minimizing a particular loss function. When a parametric family is known, maximum likelihood estimation is asymptotically efficient under certain regularity assumptions and the maximum likelihood estimators (MLEs) are used for estimating quantiles. Then, the Cramer-Rao lower bound determines the lowest possible asymptotic variance for the MLEs. Can one find a quantile estimator with a smaller variance then the Cramer-Rao lower bound? If a relevant additional information is available then the answer is yes. This manuscript considers minimum variance and mean squared error estimation which incorporate additional information for estimating optimal inventory levels.**

***Index Terms*—Newsvendor model, additional estimation, quantile estimation, minimum variance, minimum mean squared error**

I. Introduction

Newsvendor model is a popular inventory management model. This model depends on two simple quantities: a purchase price of a single unit of a product, c, and the its sale price, p. The overall profit is fully defined by the critical fractile $(= (p - c)/p)$. If F is the distribution of demand, D, then the optimal amount of product in warehouse maximizing expected profit is $Q = F^{-1}\left((p-c)/p\right)$, where F^{-1} is an inverse of F.

The distribution of the random variable D is often estimated using historic data. Hayes [1] considered exponential and Gaussian models to minimize Expected Total Operating Cost in the newsvendor problem. A Bayesian approach was used to improve estimation accuracy. Similarly, Bayesian methodology was used to solve inventory problems in [2] and [3]. In more recent literature Bayesian frameworks is also very popular.

Liyanagea and Shanthikumar [4] suggested to use direct profit maximization which simultaneously incorporates both parameter estimation and expected profit maximization.

Quantile estimation methods range from MLEs to more robust methods minimizing specific risk functions: Koenker's quantile regression [5] minimizes the sum of absolute deviations of residuals; sum of signs of residuals minimization is suggested in [6]. All of these methods are directly applicable to solving newsvendor problems as well.

Many of the above mentioned statistical methods to solving the newsvender problem lead to estimators regular enough to have two finite moments. If the two moments exist, then additional information can be combined with external information (for example, an averaged sales from another store with similar characteristics) known with a degree of uncertainty (for example, a standard error of this average maybe known) [7]. This approach assumes that the additional information is unbiased, meaning that averaged sales for both stores are about the same. A similar assumption is made in [8]. It is possible that additional information can be biased, then minimum mean squared error (MSE) can be considered instead [9], [10]. Additional information (for example, an expected value, or a quantile) known to belong to a pre-determined set of distinct values is considered in [11]–[14]; the minimum MSE criterion is also used in these papers. Zenkova and Krainova [15] consider the use of a known quantile for estimating expectations. The net premium using a known quantile for voluntary health insurance was used as an illustrative application.

This manuscript considers minimum variance and MSE estimation for incorporating additional information. Section III presents methodology for combining empirical data (historical sales data directly available for data analysis) and external information available in form of means and standard errors. Sections II and IV apply these statistical approaches for quantile estimation in newsvendor problems.

II. Illustrative Example

Table I reports an artificial dataset with 36 weeks sales data. Product A is sold at 860 dollars per unit, Product B is sold at 490 dollars per unit. A retailer pays 660 dollars/unit for Product A and 370 dollars/unit for Product B.

Product A						
6576	4263	5340	3697	3535	2651
2541	2351	3611	3867	4257	6204
6666	4364	5441	3727	3495	2755
2399	2452	3621	3961	4291	6264
6600	4333	5391	3732	3662	2498
2576	2402	3588	3900	4220	6214
Product B					
215	142	155	97	101	83
104	96	102	101	130	215
223	134	157	99	99	87
100	97	98	104	131	202
211	139	150	100	105	82
103	98	100	102	127	219

TABLE I
THIRTY SIX WEEK SALES HISTORY (IN NUMBERS OF UNITS SOLD) FOR PRODUCTS A AND B

The retailer is mainly interested in Product A as it is associated with high sales and is highly important for retailer's success. Then, the critical fractile ratio for Product A is $23.26\%(= (860-660)/860)$. The solution to the newsvendor problem is $Q = F^{-1}(0.2326)$. Relying on previous experience the retailer is confident that market demand for the two products, A and B, can be described by normal distributions, and products' demand values are likely to be correlated. Seasonal variation is so small that historic weekly data can be assumed to be independent. Normal distributions depend on two unknown parameters: μ and σ^2. Further, we will use subscripts A and B to differentiate between μ and σ^2 of Products A and B, if needed.

Using Table I, MLEs of the unknown parameters of the normal model are $\hat{\mu}_A = 4095.694$ (the sample mean) and $\hat{\sigma}_A^2 = 1791703$ (the sample variance). Note, the sample variance is slightly different from the MLE of σ_A^2, but the sample variance is an unbiased estimator of σ_A^2 and will be used instead.

Then, using the 0.2326-level normal quantile, the MLE of optimal inventory levels for Product A is $Q = 3118.14$ (in Product A units). It is possible that another estimating procedure can be chosen to evaluate Q. For example, a direct quantile estimation without any assumptions on the underlying parametric family leads to another estimate = 2859.34. For illustrative purposes, we focus on the MLE and assume that the estimate is approximately unbiased. Without loss of generality the approach on the use of additional information, considered in this paper, applies to all unbiased or approximately unbiased estimators with two finite moments.

Given the importance of Product A, it is very difficult to obtain historic sales data from similar retailers. At the same time, additional data on Product B is much simpler to obtain from other retailers. The information on Product B sales is viewed as less important by other retailers and is easily available.

Consider the following additional information. An owner of a similar retailer store claimed that his store sold 30 thousand units of Product B in past five years (additional information 1), whereas an owner of another similar store said that his sales of Product B are higher than 100 units every other week within the same five year period (additional information 2). Can these two pieces of seemingly irrelevant information be used to improve estimation accuracy of the optimal inventory levels for Product A? The answer is yes, and we will return to this illustrative example in Section IV.

III. METHODOLOGY

Let θ be a parameter of interest, $\theta = F^{-1}((p-c)/p)$ for a newsvendor problem. An estimator of θ based on historical data, $\hat{\theta}$, is assumed to be unbiased, $E(\hat{\theta}) = \theta$. In Section II, the estimator $\hat{\theta}$ is a normal quantile estimated on historical sales data ($\hat{\theta} = 3118.14$). In addition to $\hat{\theta}$, another estimator $\tilde{\eta}$ is available as additional information. This quantity estimates η not θ, and η is a different and possibly multi-dimensional parameter. In Section II, $\eta = (\eta_1, \eta_2)$, where η_1 is the mean weakly sales of Product B and η_2 is the median weakly sales of Product B. The additional information described in Section II can be converted into a two-dimensional estimate ($\tilde{\eta} = (115.3846, 100)$). The number 115.3846 is obtained as a ratio $30,000/260$, because there are 260 weeks within a five year period. The first additional information sets the mean weakly sales of Product B at 115.3846 units, and the second additional information sets the median sales at 100 units/month.

Further, we use "hat" to denote estimators based on empirical (historical) data and "tilde" for the quantities determined by additional information. Using the data in Table I, the mean weakly sales of Product B = 128 ($\hat{\eta}_1 = 128$) and weakly median sales of Product B = 103.2 ($\hat{\eta}_2$).

To combine additional information with empirical data, we consider a class of linear combinations

$$\theta^{\Lambda} = \hat{\theta} + \Lambda\left(\hat{\eta} - \tilde{\eta}\right). \tag{1}$$

In (1), $\hat{\eta}$ refers to an estimate of η based on empirical data. It is clear that $E(\hat{\eta}) = \eta$ by a property of the sample mean and median of normal data, but $E(\tilde{\eta}) = \eta + \delta$, where δ is a *possible* bias (a vector-column of biases) associated with additional information. In Section II, the bias has two components and is estimated as

$$\hat{\delta} = \hat{\eta} - \tilde{\eta} = (12.6154, 3.5).$$

Following [9], the smallest MSE in the class θ^{Λ} is secured with

$$\theta^0(\delta) = \hat{\theta} - cov\left(\hat{\theta}, \hat{\delta}\right) E^{-1}\left(\hat{\delta}\hat{\delta}^T\right)\hat{\delta}^T \tag{2}$$

and

$$MSE\left(\theta^0\right) = cov\left(\hat{\theta}\right) - cov\left(\hat{\theta}, \hat{\delta}\right) E^{-1}\left(\hat{\delta}\hat{\delta}^T\right) cov\left(\hat{\delta}, \hat{\theta}\right),$$

where $E\left(\hat{\delta}\hat{\delta}^T\right) = cov\left(\hat{\eta}\right) + cov\left(\tilde{\eta}\right) + \delta\delta^T$.

The special case of $\delta = 0$ makes θ^{Λ} unbiased for all choices of Λ. Then,

$$\theta^0(0) = \hat{\theta} - cov\left(\hat{\theta}, \hat{\delta}\right) cov^{-1}\left(\hat{\delta}\right)\hat{\delta}^T \tag{3}$$

has the smallest variance among all θ^{Λ}, see [7], and

$$cov\left(\theta^0(0)\right) = cov\left(\hat{\theta}\right) - cov\left(\hat{\theta}, \hat{\delta}\right) cov^{-1}\left(\hat{\delta}\right) cov\left(\hat{\delta}, \hat{\theta}\right).$$

For one-dimensional θ, the quadratic form on the right hand side of Equation 3

$$M = cov\left(\hat{\theta}, \hat{\delta}\right) cov^{-1}\left(\hat{\delta}\right) cov\left(\hat{\delta}, \hat{\theta}\right) \geq 0.$$

Two extreme scenarios, associated with M, describe how relevant additional information is for estimating θ:

- If $\hat{\theta}$ and $\hat{\delta}$ are uncorrelated, $M = 0$ and $\theta^0(\delta) = \hat{\theta}\ \forall \tilde{\eta}$.
- If $cov\left(\hat{\theta}, \hat{\delta}\right) = cov\left(\hat{\theta}\right)$ (exact knowledge: $\tilde{\eta} = \theta$), $M = cov\left(\hat{\theta}\right)$, $\theta^0(0) = \theta$ and $cov(\theta^0(0)) = 0$.

The estimator $\theta^0(\delta)$ is not directly applicable in practice as the covariances it relies on are unknown. In addition, the unknown δ is also present in its structure. Dmitriev and his colleagues [11] explored the same class of estimators. In contrast to our settings, they hypothesized that $\tilde{\eta} = \eta + \delta$ is known to belong to a distinct set of pre-determined values.

We estimate unknown covariances on empirical data to obtain an approximation to the optimal $\theta^0(\delta)$:

$$\hat{\theta}^0(\delta) = \hat{\theta} - \widehat{cov}\left(\hat{\theta}, \hat{\delta}\right)\left(\widehat{cov}\left(\hat{\eta}\right) + \widetilde{cov}\left(\tilde{\eta}\right) + \delta\delta^T\right)^{-1} \hat{\delta}^T. \quad (4)$$

The useful property of $\hat{\theta}^0(\delta)$ is that it is easy to show that under some regularity conditions

$$\sqrt{n}\left(\hat{\theta}^0(\delta) - \theta^0(\delta)\right) = o_p(1). \quad (5)$$

Another interesting asymptotic result is that $\forall \delta \neq 0$,

$$\sqrt{n}\left(\theta^0(\delta) - \theta\right) = o_p(1) \quad (6)$$

and

$$\sqrt{n}\left(\hat{\theta}^0(\delta) - \theta\right) = o_p(1). \quad (7)$$

From (6) and (7)

$$\sqrt{n}\left(\hat{\theta}^0(\delta) - \theta^0(\delta)\right) = o_p(1). \quad (8)$$

Estimator $\hat{\theta}^0(\delta)$, however, still includes an unknown δ. If we plug in $\hat{\delta}$ instead, we get another approximation:

$$\hat{\theta}^0(\hat{\delta}) = \hat{\theta} - \widehat{cov}\left(\hat{\theta}, \hat{\delta}\right)\left(\widehat{cov}\left(\hat{\eta}\right) + \widetilde{cov}\left(\tilde{\eta}\right) + \hat{\delta}\hat{\delta}^T\right)^{-1} \hat{\delta}^T. \quad (9)$$

The use of $\hat{\delta}$ in (4) creates certain difficulties for (5) to hold. Specifically, if $\delta = 0$ then $\sqrt{n}\left(\hat{\theta}^0(\hat{\delta}) - \theta^0(0)\right) = O_p(1)$, meaning that $\sqrt{n}\left(\hat{\theta}^0(\hat{\delta}) - \theta^0(0)\right)$ does not converge to zero, in probability, as $n \to \infty$; even asymptotically it continues to be a non-degenerate random variable.

Overall, if $\delta = 0$ can be surely assumed, minimum variance estimator $\hat{\theta}^0(0)$ is to be used, and if some protection against possible bias (disinformation) is needed minimum MSE estimation with $\hat{\theta}^0(\hat{\delta})$ is a better choice with the understanding that $\hat{\theta}^0(\hat{\delta})$ is inferior to $\hat{\theta}^0(0)$ under $\delta = 0$. The estimator $\hat{\theta}^0(\delta)$ can be used to evaluate the impact of bias on the estimating procedure.

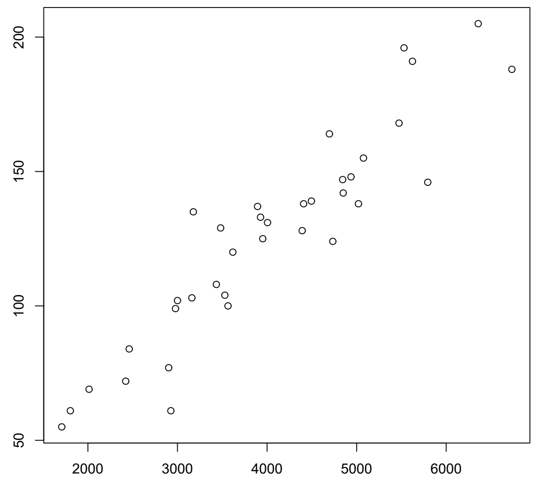

Fig. 1. Scatter plot of weekly sales of Products A (x-axis) and B (y-axis); Pearson correlation is 93.6%.

IV. Illustrative Example (continuation)

As it is shown in Section III a minimum variance estimator $\hat{\theta}^0(0)$ and a minimum MSE estimator $\hat{\theta}^0(\hat{\delta})$ are the estimators to use in practice. This section shows how both estimators can be calculated using R package “AddInf” available at

`https://github.com/starima74/AddInf`

To use “AddInf” R package one need to install “devtools” library use its “install_github” command to install “AddInf” package

```
install.packages("devtools")
library(devtools)
install_github("starima74/AddInf",
        force=TRUE)
library(AddInf)
```

The following part of R code creates weekly sales data for products A and B.

```
A <- c(6576,4263,5340,3697,3535,2651,
       2541,2351,3611,3867,4257,6204,
       6666,4364,5441,3727,3495,2755,
       2399,2452,3621,3961,4291,6264,
       6600,4333,5391,3732,3662,2498,
       2576,2402,3588,3900,4220,6214)

B <- c(215, 142, 155,97,  101, 83,
       104, 96,  102,101, 130, 215,
       223, 134, 157,99,  99, 87,
       100, 97,  98,104, 131, 202,
       211, 139, 150,100,  105, 82,
       103, 98,  100,102, 127, 219)
dd <- data.frame(A = A, B = B)
```

```
plot(dd, ylab="Product B",
     xlab="Product A",
     main="Weekly sales history")

cor(dd)
## corr = 0.936
```

First of all, distributions of sales of Products A and B need to be explored, see Figure 1. One can easily conclude that the association between the sales of Products A and B is linear and strong, which is supported by the Pearson correlation of $93.6\% (p < 0.0001)$.

The R function below estimates optimal inventory levels at the critical fractile equal to 0.2326. The data in the argument is assumed to follow a normal distribution.

```
theta.f <- function(d) {
   qnorm(0.2326, mean = mean(d$A),
   sd = sd(d$A))
}
```

A. Minimum Variance Estimation

Additional information is aggregated into a data frame in form of means ($\tilde{\eta}$) and covariances ($cov(\tilde{\eta})$). It is important to be able to estimate η on the empirical data. This is why the function implementing estimation of η needs to be defined. Here, functions are statistical procedures to calculate $\hat{\eta}$ using empirical data. We will discuss how to incorporate possible biases in Section IV-B

```
### (Empty Lists) to save
### additional Information
Add.Info.Means <- list()
Add.Info.Vars <- list()
Add.Info.Functions <- list()
Add.Info.Biases <- list()
```

The first additional data source declares weekly sales of Product B at 115.3846 units. Since $var(\tilde{\eta}_1)$ is not available as additional information, but it is stated that additional information comes from a similar store. We can realistically assume that variance is similar as well, $var(\tilde{\eta}_1) = var(\hat{\eta}_1)$. Sample variance based on 36 observations for Product B is 1912.8. Then, $var(\tilde{\eta}_1)$ is approximated by 1912.8/260. Since additional information is given by an averaged value, the function is simply the sample average.

```
Add.Info.Means[[1]] <- 115.3846
Add.Info.Vars[[1]] <- 1912.8/260
Add.Info.Functions[[1]] <-
  function(d) mean(d$B,na.rm = TRUE)
```

Information from the second source is summarized into lists in a similar manner. The second source reported median = 100. To estimate its variance, the sample on Product B is bootstrapped as follows.

```
### variance of MEDIAN
### (bootstrapping)
set.seed(123)
res <- 1:10000
for(i in 1:10000)
  res[i] <-
  median(B[sample(1:36,replace=TRUE)])
var(res)*36
### 3227.319
```

Then, 3227.319 is used to estimate variance of $\tilde{\eta}_2$ with 3227.319/260. Further,

```
Add.Info.Means[[2]] <- 100
Add.Info.Vars[[2]] <- 3227.319/260
Add.Info.Functions[[2]] <-
  function(d) median(d$B, na.rm = TRUE)
```

Note, that in the above R code η_2 is defined as the median. The lists with additional information (means, variances and functions) are aggregated into a single data frame

```
Add.Info <- data.frame(
     Means = rep(NA,2),
     Vars = rep(NA,2),
     Functions = rep(NA,2),
     Biases = rep(NA,2))
Add.Info$Means = Add.Info.Means
Add.Info$Vars = Add.Info.Vars
Add.Info$Functions = Add.Info.Functions
```

Finally, we run "MVAR" function where minimum variance with additional information is implemented. This function internally uses non-parametric bootstrap to estimate unknown covariances needed for $\hat{\theta}^0(0)$. In addition to the first three arguments described above, the function also uses number of bootstrap resamples (nboot) and a cutoff on the proportion of eigenvalues, which is a convenient way to deal with weakly definite covariance matrices.

```
res <- MVAR(dd, theta.f, Add.Info,
   nboots = 5000, eig.cutoff = 1)
```

The result consists of $\hat{\theta}^0(0)$ and its variance, and $\hat{\theta}$ and its variance.

```
res
### $`Theta.Est`
###         [,1]
### [1,] 3072.728
###
### $Theta.Est.Var
###         [,1]
### [1,] 904.6197
###
### $Theta.Hat
### [1] 3118.14
###
### $Theta.Hat.Var
### [1] 1060.981
```

From these results we can estimate the standard deviation of $\hat{\theta}^0(0)$ which is approximately equal to $30.0769(=\sqrt{904.6197})$, and the standard deviation of $\hat{\theta}$, which is $32.5727(=\sqrt{1060.981})$. Thus, the asymptotic confidence interval becomes 8% shorter with the use of additional information. Taking into account that the covariances are estimated with bootstrap resampling, the results may slightly differ from one run to another. To avoid this randomness, the "set.seed" function can be used:

```
set.seed(123)
```

The above improvement in standard deviations may seem marginal, but for other choices of additional information the changes can be more visual. For example, what if we assume that the variance of the first additional information is much smaller (for example, the 5 year sales of $30,000$ units of Product B is an averaged value across 10 different stores, which corresponds to 2600 weeks of follow-up):

```
Add.Info.Vars[[1]] <- 1912.8/2600
Add.Info$Vars = Add.Info.Vars
```

After running the MVAR function

```
res <- MVAR(dd, theta.f1, Add.Info,
   nboots = 5000, eig.cutoff = 1)
```

the result is

```
res
### $`Theta.Est`
###           [,1]
### [1,] 2962.054
###
### $Theta.Est.Var
###           [,1]
### [1,] 629.5974
###
### $Theta.Hat
### [1] 3118.14
###
### $Theta.Hat.Var
### [1] 998.8053
```

From the above, the standard deviation of $\hat{\theta}^0(0)$ is 25.09178 and the SD of $\hat{\theta}$ is 31.60388 leading to a 21% reduction in the width of the asymptotic confidence interval.

B. Minimum Mean Square Estimation

It is not impossible that the additional information came from a biased source. For example, what if the additional information claims that 5 year sales were equal to $100,000$ units of Product B, not $30,000$? This leads to a very different estimated weekly sales of 384.6154 units of Product B $(= 100,000/260)$.

```
Add.Info.Means[[1]] <- 100000/260
Add.Info$Means = Add.Info.Means
```

If this information is incorrect, it will lead to a very different and misleading optimal inventory levels for product A:

```
### $`Theta.Est`
###          [,1]
### [1,] 6238.16
###
### $Theta.Est.Var
###           [,1]
### [1,] 635.5988
###
### $Theta.Hat
### [1] 3118.14
###
### $Theta.Hat.Var
### [1] 1002.272
```

This, incorrect or deliberately altered additional information may lead to serious biases and, thus, have to be approached with extreme caution. On the other hand, if such additional information is actually correct, it can fix optimal inventory assessment based on low quality empirical data.

If, however, the empirical data is of good quality, the minimum MSE approach can be used, which provides robustness to undue influence of additional information. In this case, an indicator of possible bias needs to be added to the "Add.Info" data structure:

```
Add.Info.Biases[[1]] <- 1
Add.Info.Biases[[2]] <- 0
Add.Info$Biases = Add.Info.Biases
```

The above R code defines that the first source of additional information may be unreliable, whereas the second source is reporting unbiased additional information. In this situation, $\hat{\theta}^0(\hat{\delta})$ should be applied instead

```
res <- MMSE(dd, theta.f, Add.Info,
   nboots = 5000, eig.cutoff = 1)
res
###  $`Theta.Est`
###           [,1]
### [1,] 3110.567
###
### $Theta.Est.Var
###           [,1]
### [1,] 976.0103
###
### $Theta.Hat
### [1] 3118.14
###
### $Theta.Hat.Var
### [1] 1045.683
```

Now, the $\hat{\theta}^0(\hat{\delta}) = 3110.567$ and is very close to $\hat{\theta} = 3118.14$. The unreliable additional information is automatically suppressed, as its serious bias is easily detected, whereas the second source still contributed to a smaller asymptotic variance 976.0103 of $\hat{\theta}^0(\hat{\delta})$ versus 1045.683 of $\hat{\theta}$.

If the bias is small, for example ($\tilde{\eta}_1 = 130$ and $\hat{\eta}_1 = 128$)

```
Add.Info.Means[[1]] <- 130
Add.Info$Means = Add.Info.Means
```

then

```
res <- MMSE(dd, theta.f, Add.Info,
    nboots = 5000, eig.cutoff = 1)
res
### $`Theta.Est`
###         [,1]
### [1,] 3114.023
###
### $Theta.Est.Var
###         [,1]
### [1,] 647.4465
###
### $Theta.Hat
### [1] 3118.14
###
### $Theta.Hat.Var
### [1] 1002.441
```

Thus, when additional information is consistent with empirical data, MMSE and minimum variance approaches show similar improvement of variance.

V. SUMMARY

Additional information in form of known statistical quantities and their standard errors can be helpful in estimating many statistical quantities including optimal inventory levels in newsvendor models. This manuscript shows how to incorporate such additional information into statistical estimation. An illustrative example on estimating optimal inventory levels with additional information is analyzed with the R package "AddInf".

The illustrative example shows how information from multiple additional sources can be used. If additional information is not correct or deliberately altered (disinformation) the minimum variance estimation may be inappropriate. At the same time minimum mean squared estimation detects that such additional information is inconsistent with the empirical data and its impact is suppressed. On the other hand if the additional information is consistent with the empirical data, the benefits of using minimum variance and minimum MSE approach are comparable.

Thus, with the use of additional information, more accurate assessment of optimal inventory levels is obtained.

REFERENCES

[1] R. H. Hayes, "Statistical estimation problems in inventory control," *Manag. Sci.*, vol. 15, no. 11, pp. 686–701, 1969.

[2] H. Scarf, "Bayes solutions of the statistical inventory problem," *Ann. of Math. Stat.*, vol. 30, no. 2, pp. 490–508, 1959. [Online]. Available: http://www.jstor.org/stable/2237094

[3] S. Karlin, "Dynamic inventory policy with varying stochastic demands," *Manag. Sci.*, vol. 6, no. 3, pp. 231–258, 1960. [Online]. Available: http://www.jstor.org/stable/2627339

[4] L. H. Liyanagea and J. Shanthikumar, "Apractical inventory control policy using operational statistics," *Oper. Res. Lett.*, no. 33, pp. 341–348, 2005.

[5] R. Koenker, *Quantile Regression*, ser. Econometric Society Monographs. Cambridge University Press, 2005.

[6] F. Tarassenko, S. Tarima, A. Zhuravlev, and S. Singh, "On sign-based regression quantiles," *J. Stat. Comput. Simul.*, vol. 85, no. 7, pp. 1420–1441, 2015.

[7] S. Tarima and D. Pavlov, "Using auxiliary information in statistical function estimation," *ESAIM: Probab. Stat.*, vol. 10, pp. 11–23, 2006.

[8] S. Tarima, S. Slavova, T. Fritsch, and L. Hall, "Probability estimation when some observations are grouped," *Stat. Med.*, vol. 26, no. 8, pp. 1745–1761, 2007.

[9] S. Tarima and Y. Dmitriev, "Statistical estimation with possibly incorrect model assumptions," *Bul. Tomsk St. University: cont., comput., inf.*, vol. 8, pp. 78–99, 2009.

[10] S. Tarima, A. Vexler, and S. Singh, "Robust mean estimation under a possibly incorrect log-normality assumption," *Commun. Stat.–Simul. C.*, vol. 42, no. 2, pp. 316–326, 2013.

[11] Y. Dmitriev, P. Tarassenko, and Y. Ustinov, "On estimation of linear functional by utilizing a prior guess," in *Information Technologies and Mathematical Modelling*, A. Dudin, A. Nazarov, R. Yakupov, and A. Gortsev, Eds. Cham: Springer International Publishing, 2014, pp. 82–90.

[12] Y. Dmitriev and P. Tarassenko, "On adaptive estimation using a prior guess," in *Applied methods of statistical analysis. Nonparametric approach - AMSA2015*, Novosibirsk, September, 2015, pp. 49–55.

[13] Y. Dmitriev, P. Tarassenko, and F. Tarassenko, "On improving statistical estimation by utilizing collateral information (guesses): a case of the probability estimation," in *International Workshop on Applied Methods of Statistical Analysis: Nonparametric methods in cybernetics and system analysis*, Krasnoyarsk, September, 2017, pp. 262–269.

[14] Y. Dmitriev, G. Koshkin, and V. Lukov, "Combined identification and prediction algorithms," in *IV International Research Conference: Information Technologies in Science, Management, Social Sphere and Medicine*, Tomsk, December, 2017, pp. 244–247.

[15] Z. Zenkova and E. Krainova, "Estimating the net premium using additional information about a quantile of the cumulative distribution function," *Bus. Inform.*, vol. 42, no. 4, pp. 55–63, 2017.