Journal of Statistics Applications & Probability
An International Journal

© 2012 NSP
Natural Sciences Publishing Cor.

Statistical Aspects of Forest Harvesting: Price -Weighted Apportionment Index and Related Inference

J V Sukanya Divvela[1] *and Bikas K Sinha*[2]

[1] JRF(Statistics), C R Rao Advanced Institute of Mathematics, Statistics and Computer Science, Hyderabad
[2] Indian Statistical Institute, Kolkata

Received: Dec. 12, 2011; Accepted Feb. 10, 2012
Published online: 1 April 2012

Abstract: Price-Weighted Apportionment Index measures the fit between *log* demand distribution and *log* output distribution. We present the asymptotic sampling distribution of Price-Weighted Apportionment Index by assuming a multinomial distribution for the outcome variables. Our results are based mainly on large-sample normal approximation.

Keywords: Target Matrix, Output Matrix, Price Matrix, Apportionment Index(AI), Price-Weighted Version of Apportionment Index(AI^w), Lower Tolerance Limit.

1. Introduction

This paper is based on statistical considerations of topics related to *log bucking outcome in cut-to-length (CTL) forest harvesting.* The first steps towards a fully mechanized forest harvesting industry were taken about 50 years ago when the first forest harvesters, i.e. forest machines capable of felling, delimbing and bucking trees, were introduced (Drushka & Konttinen, 1997; Gellerstedt & Dahlin, 1999). The degree of mechanization, however, varies considerably between different countries. In the Nordic countries, for example, almost all harvesting is currently done mechanically, while in many Eastern European countries the traditional motor-manual methods still dominate (Axelsson, 1998; Asikainen et al., 2005). According to a rough estimate (Ponsse Oyj, 2006), about 45% of the world's annual cutting volume is currently harvested mechanically. The degree of mechanization, however, is expected to further increase worldwide as the forestry industry focuses on reducing costs, improving productivity and concentrating on labor-related issues (Murphy, 2002).

Mechanized harvesting can be divided into three main methods which differ in terms of the amount of processing done at the harvesting site in the forest (Pulkki, 1997; Owende, 2004). (1) In the cut-to-length method trees are felled, delimbed and bucked into shorter logs directly upon felling. The resulting logs are then transported by a forwarder to the roadside and further by timber truck to the production plant(s) for further processing. (2) In the tree-length method (TL) trees are only topped (i.e. the top of a tree is cut off at a pre-determined minimum diameter) and delimbed in the forest. The 5 bucking is done at the separate terminal or at the mill's log yards. (3) In the whole tree method (also known as the full tree method) trees are felled and forwarded to the roadside with branches and top intact. The whole (full) trees are further processed either at the roadside or, after haulage, at the central processing yard or the mill.

Although the popularity of the CTL method is steadily growing, it still today accounts for less than half of the world's roundwood harvest (Asikainen et al., 2005). A rough estimate of its current share in the world's mechanically harvested timber is about 35% (Ponsse Oyj, 2006). In Finland and Sweden almost all harvesting is carried out by CTL systems (Gellerstedt & Dahlin, 1999). The CTL method is also re-establishing itself in North America, where the TL and full tree systems have traditionally been the dominant harvesting methods (Pulkki, 1997).

* Corresponding author: e-mail:

© 2012 NSP
Natural Sciences Publishing Cor.

Most harvesters currently employed in CTL operations are single-grip models. A single-grip harvester has only one unit for both felling and reproducing processes mounted on an articulating arm. A double-grip (two-grip) harvester, which was popular in the 1970s, has two separate units; one for felling and the other for the delimbing, bucking and sorting processes.

The first CTL harvesters with automatic measuring systems came onto the market in the early 1970s. These first measuring systems, however, could measure and record only tree length. The capability of continuously measuring tree diameter while harvesting was not incorporated into them until the mid 1980s (Marshall, 2005; Drushka & Konttinen, 1997). Today harvesters are equipped with high-class information systems able not only to measure the dimensions of trees but also to predict the stem profile of each tree being processed and thereby to tailor the bucking outcome for the desired output. They have thus become an important part of the logistics chain from the forest to the end user. To optimize the overall flow, more recent development has focused on utilizing modern information technology such as geographical and positioning systems (GIS and GPS), online internet applications and information transfer over mobile phones.

In the course of processing, the harvester first fells the tree and then runs it through the processing unit (i.e., a harvester head in a single-grip or a delimbing-cutting device in a double-grip harvester). The length along the stem is simultaneouly measured either by the running wheel located at the harvester head (90% of all heads) or on the feed-rollers (Gellerstedt, 2002). The stem diameter is usually measured by the amount of opening in the delimbing knives or the feed-rollers using a cross measure. In measuring the stem, the data are simultaneously stored in an on-board computer. Before starting bucking optimization, filtering or smoothing techniques are used to eliminate the most crucial discrepancies in the measured data.

The common trend in the sawmill industry, at least in Scandinavia, is towards customer-oriented production of well-defined products. In fact, controlling the wood flow from forest to mills in such a way that the mills' requirements are satisfied has recently been seen as an even more important development area in wood procurement than the traditional attempt to reduce transportation and harvesting costs (Kivinen, 2006). As customer-oriented production strategies have gained ground in the sawmill industry, it has become more and more important not only to supply the sawmill with a sufficient number of logs at minimum cost, but also to ensure that the raw material meets the requirements of the sawmill as regards length, diameter and quality distribution of logs (Kivinen, 2004). This, in turn, has made proper assessment of the goodness of the bucking outcome of crucial importance.

In general, there are two situations where the agreement between the distribution of logs demanded by the sawmill (demand distribution) and the actual outcome (output) distribution of logs is of particular interest. These are (1) the standard pre-harvest planning procedure where most suitable stands for prevailing customer orders need to be determined, and (2) the postharvest analysis where it may be desirable to know, for example, how various harvesters have succeeded in meeting a certain demand distribution or to determine whether there are any significant differences between various wood suppliers. A proper measure for evaluating the bucking outcome also provides information on how to adjust the bucking instructions to meet the desired log distribution (Kivinen et al., 2005).

We organize the rest of this paper as follows. In Section 2, following Nummi et al (2005), Sinha et al (2005a) and Sinha et al (2005b), among others, we introduce and breifly reveiw the concept of 'Apportionment Index' and related formulae for its measurement. In Section 3, we discuss the main distributional properties of Price-Weighted Apportionment Index(AI^w). In Section 4, we give the statistical analysis of AI^w along with an illustrative example. We consider the problems of (1) determination of sample size N for the AI^w to attain a preassigned lower tolerance limit with a specified confidence level, and (2) determination of a lower tolerance limit to be attained by the AI^w for given sample size N, again with a specified confidence level. In section 5, we make some concluding remarks and discuss scope for further research.

2. Measuring the Bucking Outcome

2.1. Target, Outcome and Price Matrix

The outcome of the actual harvesting operation has been measured mainly by comparing the relative proportions of the output and target distributions. More specifically, let

$$\theta = (\theta_{ij}) = \theta_{11}\theta_{12}...............\theta_{1n} \quad \theta_{21}\theta_{22}...............\theta_{2n} \quad ...... \quad ...... \quad ...... \quad \theta_{m1}\theta_{m2}...............\theta_{mn} \qquad (1)$$

denote the $m \times n$ demand (target) matrix for a certain log type, where each row represents a particular small end diameter (SED) class of logs, each column refers to a particular length class and t_{ij} is the number of logs in the i^{th} diameter class and j^{th} length class, $i = 1, ..., m$ and $j = 1, ..., n$. A log with an SED of d and a length of l will belong to the log class (i, j) if the log satisfies the constraints $d_i\, d < d_{i+1} and\, l_j,\, l < l_{j+1}$. Correspondingly, $m \times n$ matrix

$$X = (X_{ij}) = X_{11}X_{12}...............X_{1n} \quad X_{21}X_{22}...............X_{2n} \quad ...... \quad ...... \quad ...... \quad X_{m1}X_{m2}...............X_{mn}. \qquad (2)$$

© 2012 NSP
Natural Sciences Publishing Cor.

denotes the outcome of the harvesting operation. The $m \times n$ price matrix specifies relative prices for all log categories, i.e. determines how valuable or profitable it is to cut different length-diameter combinations of a particular log type. The price matrix can be given as

$$P = (p^*_{ij}) = p^*_{11}p^*_{12}..............p^*_{1n} \quad p^*_{21}p^*_{22}..............p^*_{2n} \quad ...... \quad ...... \quad ...... \quad p^*_{m1}p^*_{m2}..............p^*_{mn}. \quad (3)$$

where $p^*_{ij} = \frac{p_{ij}}{\sum\sum p_{ij}}$ is the relative price of the i^{th} diameter and j^{th} length combination of logs and p_{ij} is the respective absolute price.

2.2. Some Measures for Evaluating the Log Bucking Outcome

A common practice in Scandinavia is to evaluate the fit between the demand and actual output log distributions with the Apportionment Index (AI) or Apportionment Degree, first introduced in forestry by Bergstrand in the mid- 1980s (e.g. Bergstrand, 1989). For a fixed quality class the AI is defined as

$$AI = 1 - \frac{1}{2}\sum_{i=1}^{m}\sum_{j}^{n}|X^*_{ij} - \theta^*_{ij}| \quad (4)$$

where $X^*_{ij} = \frac{X_{ij}}{\sum\sum X_{ij}}$ and $\theta_{ij} = \frac{\theta_{ij}}{\sum\sum \theta_{ij}}$ are the relative proportions of the outcome and target matrices, respectively. After some simple manipulations it can be shown that the AI can be rewritten as

$$AI = \sum_{i=1}^{m}\sum_{j}^{n} min(X_{ij}, \theta_{ij}). \quad (5)$$

The maximum value of the AI is 1 (100%), which indicates a perfect match between the distributions. The minimum value of the index is $min(\theta_{11}, \theta_{12}, ..., \theta_{mn})$, i.e. the smallest relative cell target, which is reached when all the logs fall into the diameter-length class of the smallest target proportion. In some of the original papers this kind of a scenario is referred to as a perfect mismatch.

The AI may be interpreted as the proportion of the ”correctly” located logs in the outcome distribution with respect to the demanded log distribution. For example, if the AI value were 0.85, this would mean that 85% of the produced logs are in accordance with the demanded distribution while 15% are of the wrong size and should have been allocated to other log categories during the bucking process to make the outcome equal to the target, i.e. to attain complete agreement between the two distributions. In fact, by observing the deviation of the outcome from the target matrix in terms of upload or download, i.e. $c_{ij} = X_{ij} - \theta_{ij}$, the AI can also be expressed for equal matrix totals as

$$AI = \frac{N - \sum_{i=1}^{m}\sum_{j}^{n} c_{ij} I(c_{ij} > 0)}{N}. \quad (6)$$

where $N = \sum_{i=1}^{m}\sum_{j}^{n} X_{ij} = \sum_{i=1}^{m}\sum_{j}^{n} \theta_{ij}$ and $I(c_{ij} > 0) = 1$ for $c_{ij} > 0$ and 0 otherwise.

The AI has gained ground especially by merit of its simplicity, easy interpretability and ease of use. The measure has been criticized mainly as being too crude, since, for example, it attributes the same weight to all log classes. Hence, a price-weighted version of the AI was proposed by Kivinen et al.(2005) and Nummi et al. (2005). The price-weighted Apportionment Index utilizes the price matrix and is defined as

$$AI^w = \sum_{i=1}^{m}\sum_{j=1}^{n} p^*_{ij} min(X^*_{ij}, \theta^*_{ij}) \quad (7)$$

The AI^w is not as amenable to interpretation as the non-weighted AI, which is clearly seen as a disadvantage of the measure.

Some penalty-based variants of the traditional AI were proposed in Kirkkala et al.(2000), Weijo (2000) and Malinen & Palander (2004). The idea of using prices as weights when measuring the agreement of the two distributions lead Kivinen et al. (2005) to apply the theory of index numbers common in economics. The authors suggested the use of the Laspeyres' quantity index to describe the relationship between the values of the postharvest and preharvest log distributions. We will not discuss these incidences in our paper.

Instead of using the Apportionment Index or its derivatives to evaluate the similarity between the demand and output log distributions, standard statistical tests can also be applied. The most commonly used test for examining the goodness-of-fit of grouped data is the frequency $\chi^2 - test$, which was applied in the forestry context e.g. in Malinen & Palander

© 2012 NSP
Natural Sciences Publishing Cor.

(2004), Kivinen et al. (2005) and Nummi et al. (2005). Using the same notations as above, the test statistic can be defined as

$$\chi^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} \frac{(X_{ij} - \theta_{ij})^2}{\theta_{ij}}. \tag{8}$$

In the case of a perfect match the value of the χ^2-statistic equals zero. However, as the deviation between the two matrices increases, the value of the measure also increases, giving large positive values for large deviations. Kivinen et al. (2005) solved the scaling problem of the $\chi^2 - statistic$ by using the contingency coefficient C defined as

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}}. \tag{9}$$

where N is the total number of logs harvested. Subtracting the contingency coefficient from 1 then yields a measure which equals 1 for perfect match and tends to decrease towards 0 as the deviation between the distributions increases. Nummi et al. (2005), however, solved the scaling problem by utilizing the p-value assigned to the χ^2-statistic. The AI is closely related to e.g the Dissimilarity Index (DI) or Index of Dissimilarity commonly used in sociology for measuring segregation. One of the very first instances of the DI as a measure of segregation was that in the paper by Jahn et al. (1947). The DI is also commonly used to summarize the closeness of fit of a model to the categorical sample data (e.g. Agresti, 2002, pp. 329-330). The so-called overlapping coefficient (OVL) was later defined as a generalized measure of agreement or similarity between two probability distributions or two populations represented by such distributions (Inman & Bradley, 1989). If $f_1(x)$ and $f_2(x)$ are density functions defined on the n-dimensional Euclidian space R_n, then the OVL can be defined as

$$OVL = \int_{R_n} min[f_1(x), f_2(x)]dx \tag{10}$$

In a simple univariate case the OVL is simply the fraction of the probability mass common to both distributions. In a case of two discrete probability distributions, the relation of the OVL to the Apportionment Index and Dissimilarity Index can be expressed as $OVL = 1 - DI = AI$.

Although the traditional AI is today the measure most widely used for assessing the agreement between the demand and the output log distributions, its superiority over the other measures is somewhat questionable. It is not easy to make comparisons between the measures, since, first, they differ in scaling and, second, there exists no commonly approved yardstick capable of giving the ”true” ranking of all possible bucking outcomes with respect to the given demand distribution. Kivinen et al. (2005) approached the problem of comparing different measures by defining four criteria for an ideal measure. Four alternative goodness-of-fit measures were then tested against the criteria. The tested measures were: (1) the traditional AI, (2) the $\chi^2 - statistic$, (3) the Laspeyres' quantity index and (4) the price-weighted AI. The results of the study showed no marked differences between the performances of the four measures compared. Neither did the results indicate the universal superiority of any of the candidates. All four measures met three of the four requirements of an ideal measure and provided fairly consistent results for different demand matrices in different stand types. Malinen & Palander (2004) compared the performance of five alternative goodness-of-fit measures on the basis of their ability to control the bucking-to-demand procedure. Since the use of the goodness-of-fit measures in the online control of the bucking procedure is not a topic of this paper, we may content ourselves with the above reference to this particular study.

3. AI^w- Distributional Results

We will use the notations and terminologies as in the above, mostly without explicit mention of the source of the problem such as 'log distribution' of different 'width-length dimensions'. Our results are very general and apply in similar contexts. An Appendix at the end includes some basic results used for the mathematical derivations in this section. We assume that the observed frequency counts in the given matrix $x = ((x_{ij}))$ of dimension $m \times n$ in different cross-classified cells constitute a realization of random frequency counts in the matrix $X = ((X_{ij}))$ of the same dimension. Further, $((X_{ij}))$'s follow the standard multinomial distribution i.e. $(X_{11}, X_{21} \ldots\ldots X_{mn}) \sim MN(N; \theta^*_{11}, \theta^*_{21} \ldots\ldots \theta^*_{mn})$.
where θ^*_{ij}'s are the cell probabilities with $\theta^*_{ij} \in (0, 1)$ and $\sum\sum \theta^*_{ij} = 1$.

Remark 1 Again, from forestry application perspective, the distributional assumption may not necessarily make sense if a stand is tried to harvest by a target which is unsuitable for the particular stand. It is, for example, impossible to produce 'logs' with large Small End Diameter from trees with smaller Diameter at Breast Height. A situation of this kind immediately leads to a poor fit between the outcome and demand distribution arguing that the outcome cannot be a realization

© 2012 NSP
Natural Sciences Publishing Cor.

from the given multinomial distribution. However, if the pre-harvest planning procedure is performing appropriately, it seems reasonable that the selection of the harvested stands is done such that the production of the requested logs is possible.

Once more, we reiterate that for the relative random outcome, we have used the notation $X^*_{ij} = (X^*_{11}, X^*_{21} \ldots\ldots X^*_{mn})$, where $X^*_{ij} = \frac{X_{ij}}{N} \in (0,1)$ and $\sum\sum X^*_{ij} = 1$. It readily follows that $X_{ij} \sim Bin(N, \theta^*_{ij})$ and hence,

$$E(X^*_{ij}) = E\left(\frac{X_{ij}}{N}\right) = \theta^*_{ij}. \tag{11}$$

$$V\left(X^*_{ij}\right) = V\left(\frac{X_{ij}}{N}\right) = \frac{\theta^*_{ij}\left(1-\theta^*_{ij}\right)}{N} = \sigma^2_{ij}. \tag{12}$$

$$Cov\left(X^*_{ij}, X^*_{rs}\right) = -\frac{\theta^*_{ij}\theta^*_{rs}}{N}, (ij) \neq (rs). \tag{13}$$

$$Corr\left(X^*_{ij}, X^*_{rs}\right) = -\sqrt{\frac{\theta^*_{ij}\theta^*_{rs}}{\left(1-\theta^*_{ij}\right)\left(1-\theta^*_{rs}\right)}} = \rho_{ij,rs}. \tag{14}$$

Under certain conditions, the binomial distribution can be approximated by the normal distribution. A conservative rule to follow in the case of $X \sim Bin(n,p)$ is that $min(np, n(1-p)) \geq 5$ (see e.g. Casella and Berger 2002, p. 104-105). The normal approximation gives $\frac{\sqrt{N}\left(X^*_{ij}-\theta^*_{ij}\right)}{\sqrt{\theta^*_{ij}\left(1-\theta^*_{ij}\right)}} \sim N(0,1)$.

If we assume normal distribution, upon application of Result A.1 in Appendix, $E(AI^w)$ becomes

$$E(AI^w) = E\left(\sum_{i=1}^m \sum_{j=1}^n p^*_{ij}\left[\frac{X^*_{ij}+\theta^*_{ij}}{2} - \frac{\left|X^*_{ij}-\theta^*_{ij}\right|}{2}\right]\right)$$
$$= \sum_{i=1}^m \sum_{j=1}^n p^*_{ij} E\left[\frac{X^*_{ij}+\theta^*_{ij}}{2} - \frac{\left|X^*_{ij}-\theta^*_{ij}\right|}{2}\right]$$
$$= \frac{1}{2N}\sum_{i=1}^m \sum_{j=1}^n p^*_{ij}\left(N\theta^*_{ij} + N\theta^*_{ij}\right) - \frac{1}{2N}\sum_{i=1}^m \sum_{j=1}^n p^*_{ij}\sigma_{ij}\sqrt{\frac{2}{\pi}}$$
$$= \sum_{i=1}^m \sum_{j=1}^n p^*_{ij}\theta^*_{ij} - \frac{1}{\sqrt{2N\pi}}\sum_{i=1}^m \sum_{j=1}^n p^*_{ij}\sqrt{\theta^*_{ij}\left(1-\theta^*_{ij}\right)}.$$
$$= \overline{\theta}^*(P^*) - \frac{D(\theta)}{\sqrt{N}}.$$

where

$$D\left(\theta\right) = \frac{1}{\sqrt{2\pi}}\sum_{i=1}^m \sum_{j=1}^n p^*_{ij}\sqrt{\theta^*_{ij}\left(1-\theta^*_{ij}\right)}. \tag{15}$$

It follows from (1) that $AI^w \leq \sum\sum p^*_{ij}\theta^*_{ij} = \overline{\theta}^*(P^*)$, say, with probability 1. Hence, trivially, $E(AI^w) \leq \overline{\theta}^*(P^*)$, which is also evident from the above. Further, it is seen that as N increases, the limiting value of $E(AI^w)$ coincides with $\overline{\theta}^*(P^*)$.

We now proceed to compute an expression for the variance of AI^w. There are technical details which are relegated to the appendix. The main results are listed below.

$$\text{V}(AI^w) = V\left(\frac{1}{2}\sum_{i=1}^m \sum_{j=1}^n p^*_{ij}\theta^*_{ij} + \frac{1}{2N}\sum_{i=1}^m \sum_{j=1}^n p^*_{ij}\left(X_{ij} - |X_{ij}-\theta_{ij}|\right)\right)$$
$$= \frac{1}{4N^2}V\left(\sum_{i=1}^m \sum_{j=1}^n p^*_{ij}\left(X_{ij} - |X_{ij}-\theta_{ij}|\right)\right)$$
$$= \frac{1}{4N^2}\left[\sum_{i=1}^m \sum_{j=1}^n V\left(p^*_{ij}\left(X_{ij} - |X_{ij}-\theta_{ij}|\right)\right)\right]$$
$$+ \frac{1}{4N^2}Cov\left[\left(\sum_{i=1}^m \sum_{j=1}^n p^*_{ij}\left(X_{ij} - |X_{ij}-\theta_{ij}|\right), \sum_{r=1}^m \sum_{s=1,(ij)\neq(rs)}^n p^*_{rs}\left(X_{rs} - |X_{rs}-\theta_{rs}|\right)\right)\right]$$
$$= \frac{1}{4N^2}\left[\sum_{i=1}^m \sum_{j=1}^n (p^*_{ij})^2 V\left(X_{ij} - |X_{ij}-\theta_{ij}|\right)\right]$$
$$+ \frac{1}{4N^2}\left[Cov\left(\sum_{i=1}^m \sum_{j=1}^n p^*_{ij}X_{ij}, \sum_{r=1}^m \sum_{s=1,(ij)\neq(rs)}^n p^*_{rs}X_{rs}\right)\right]$$
$$+ \frac{1}{4N^2}Cov\left(\sum_{i=1}^m \sum_{j=1}^n p^*_{ij}\left|X_{ij}-\theta_{ij}\right|, \sum_{r=1}^m \sum_{s=1,(ij)\neq(rs)}^n p^*_{rs}\left|X_{rs}-\theta_{rs}\right|\right)$$
$$- \frac{1}{4N^2}Cov\left(\sum_{i=1}^m \sum_{j=1}^n p^*_{ij}X_{ij}, \sum_{r=1}^m \sum_{s=1,(ij)\neq(rs)}^n p^*_{rs}\left|X_{rs}-\theta_{rs}\right|\right)$$
$$- \frac{1}{4N^2}Cov\left(\sum_{i=1}^m \sum_{j=1}^n p^*_{ij}\left|X_{ij}-\theta_{ij}\right|, \sum_{r=1}^m \sum_{s=1,(ij)\neq(rs)}^n p^*_{rs}X_{rs}\right)$$

© 2012 NSP
Natural Sciences Publishing Cor.

$$= \frac{1}{4N^2}\left[\sum_{i=1}^{m}\sum_{j=1}^{n}2(p^*_{ij})^2\left(N\theta^*_{ij}\left(1-\theta^*_{ij}\right)\right)(1-\frac{1}{\pi})+(A+B-C-D)\right]$$
$$= \frac{1}{4N^2}\left[2\sum_{i=1}^{m}\sum_{j=1}^{n}(p^*_{ij})^2\left(N\theta^*_{ij}\left(1-\theta^*_{ij}\right)\right)(1-\frac{1}{\pi})\right]$$
$$-\frac{1}{4N^2}\sum\sum_{(i,j)}\sum\sum_{(r,s),(i,j)\neq(r,s)} p^*_{ij}p^*_{rs}N\theta^*_{ij}\theta^*_{rs}$$
$$+\frac{1}{2\pi N^2}\sum\sum_{(i,j)}\sum\sum_{(r,s),(i,j)\neq(r,s)} p^*_{ij}p^*_{rs}\sqrt{N\theta^*_{ij}\left(1-\theta^*_{ij}\right)}\sqrt{N\theta^*_{rs}\left(1-\theta^*_{rs}\right)}$$
$$\left[\rho_{ij,rs}sin^{-1}\rho_{ij,rs}+\sqrt{\left(1-\rho^2_{ij,rs}\right)}-1\right]$$
$$= \frac{V\left(\theta\right)}{N}.$$

$$V\left(AI^w\right)=\frac{V\left(\theta\right)}{N}. \quad (16)$$

where

$$V\left(\theta\right)=\frac{1}{2}(1-\frac{1}{\pi})\left[\sum_{i=1}^{m}\sum_{j=1}^{n}(p^*_{ij})^2\left(\theta^*_{ij}\left(1-\theta^*_{ij}\right)\right)\right]-\frac{1}{4}\sum\sum_{(i,j)}\sum\sum_{(r,s),(i,j)\neq(r,s)}p^*_{ij}p^*_{rs}\theta^*_{ij}\theta^*_{rs}$$
$$+\frac{1}{2\pi}\sum\sum_{(i,j)}\sum\sum_{(r,s),(i,j)\neq(r,s)}p^*_{ij}p^*_{rs}\sqrt{\theta^*_{ij}\left(1-\theta^*_{ij}\right)}\sqrt{\theta^*_{rs}\left(1-\theta^*_{rs}\right)}$$
$$\left[\rho_{ij,rs}sin^{-1}\rho_{ij,rs}+\sqrt{\left(1-\rho^2_{ij,rs}\right)}-1\right].$$

In the above,
(i) 'A' involves $Cov\left(X^*_{ij},X^*_{rs}\right)=-\frac{\theta^*_{ij}\theta^*_{rs}}{N},\ (ij)\neq(rs)$.
(ii) 'B' involves $Cov\left(\left|X-\theta_1\right|,\left|Y-\theta_2\right|\right)$ and this has been worked out in the Results A.4 [expression (13)] of the appendix.
(iii) 'C' and 'D' involve the expressions of the form $Cov\left(X_{ij},\left|X_{rs}-\theta_{rs}\right|\right)$ which are shown to be zero in the Result A.3 of the appendix.

4. Statistical Analysis of Price-Weighted Apportionment Index

4.1. Illustrative Example

The following are the target matrix, output matrix and the price matrix taken from Sinha et al(2005b).

$$\theta^*=\begin{bmatrix}0.028 & 0.028 & 0.058 & 0.045 & 0.045\\ 0.037 & 0.017 & 0.065 & 0.045 & 0.037\\ 0.017 & 0.049 & 0.037 & 0.044 & 0.055\\ 0.022 & 0.039 & 0.039 & 0.044 & 0.059\\ 0.019 & 0.030 & 0.047 & 0.054 & 0.052\end{bmatrix}\quad x^*=\begin{bmatrix}0.02889 & 0.01444 & 0.06222 & 0.04111 & 0.04556\\ 0.03778 & 0.006667 & 0.06889 & 0.04667 & 0.03556\\ 0.01667 & 0.045556 & 0.03444 & 0.04556 & 0.05889\\ 0.02222 & 0.04 & 0.04 & 0.04444 & 0.06333\\ 0.01778 & 0.02667 & 0.04778 & 0.05444 & 0.05444\end{bmatrix}$$

$$P^*=\begin{bmatrix}0.02785 & 0.02868 & 0.02924 & 0.03008 & 0.03035\\ 0.03453 & 0.03564 & 0.03620 & 0.03732 & 0.03760\\ 0.04010 & 0.04121 & 0.04205 & 0.04344 & 0.04372\\ 0.04344 & 0.04483 & 0.04567 & 0.04678 & 0.04734\\ 0.04456 & 0.04595 & 0.04678 & 0.04818 & 0.04845\end{bmatrix}$$

The price weighted observed apportionment index $AI^w=0.0394$. The maximum value of AI^w i.e., $\overline{\theta}^*(P^*)$ for the given dataset is 0.04042. The standardized value of AI^w is $\frac{AI^w}{\overline{\theta}^*(P^*)}=0.9739$.Moreover $D\left(\theta\right)=0.0772$ and $V\left(\theta\right)=0.00016$.

© 2012 NSP
Natural Sciences Publishing Cor.

Table 1:Display of sample size N for a pre-assigned lower tolerence limit $\gamma\left(\overline{\theta}^*(P^*)\right)$ to the AI^w, with a confidence limit of 95% / 99%.

2*γ	2***lower tolerence limit** = $\gamma\left(\overline{\theta}^*(P^*)\right)$	$\alpha = 0.05$	α= 0.01
		N	N
0.5	2.02%	24	28
0.6	2.43 %	37	44
0.7	2.83 %	66	78
0.8	3.23 %	147	175
0.9	3.64%	590	699
0.95	3.84%	2358	2796

4.2. Statistical Analysis

For the output and target matrices to be considered in agreement with each other (keeping in veiw the importance of the price matrix as well), the lower tolerance limit serves as the least acceptable value of the AI^w. Thus, in a case where the observed outcome matrix assigns the index value which falls below the lower tolerance limit, we would consider the outcome to indicate disagreement with the given target matrix. Overall, the larger the observed AI^w, the better the agreement. We hence suggest two approaches to indicate a satisfactory level of agreement:

(1) The first approach uses the upper limit $\left(\overline{\theta}^*(P^*)\right)$ of AI^w or of $E(AI^w)$ as the point of reference;

(2) The second approach uses the mean$(E(AI^w))$ as the point of reference.

It is worth noting, especially for practical implementation, that $E(AI^w)$ is very easy to compute.

The first approach is as follows :

$P\left(AI^w > \gamma\overline{\theta}^*(P^*)\right) \geq 1-\alpha$

$$\Rightarrow Z_{1-\alpha} = \frac{\gamma\overline{\theta}^*(P^*) - E(AI^w)}{\sqrt{V\left(AI^w\right)}}. \tag{17}$$

Recall that $E(AI^w) = \overline{\theta}^*(P^*) - \frac{D(\theta)}{\sqrt{N}}$ while $V\left(AI^w\right) = \frac{V(\theta)}{N}$. From (9), we can solve for N and γ interchangeably for given α. Further note that $Z_{1-\alpha} = -Z_\alpha$. Specifically we have

Formula for N:

$$N = \left[\frac{D\left(\theta\right) + Z_\alpha\sqrt{V\left(\theta\right)}}{\overline{\theta}^*(P^*)\left(1-\gamma\right)}\right]^2. \tag{18}$$

Formula for γ:

$$\gamma = 1 - \left[\frac{D\left(\theta\right) + Z_\alpha\sqrt{V\left(\theta\right)}}{\overline{\theta}^*(P^*)\sqrt{N}}\right]. \tag{19}$$

In the tables below we provide some numerical computation with reference to the target matrix and the price matrix displayed above. Note that $\overline{\theta}^*(P^*) = 4.04\%$.

The second approach involving $E(AI^w)$ as a reference point, is as follows:

$P\left(AI^w > E(AI^w)\left(1-\epsilon\right)\right) \geq 1-\alpha,$

$\Rightarrow Z_{1-\alpha} = -Z_\alpha = -\frac{E(AI^w)\epsilon}{\sqrt{V(AI^w)}}.$

For a given α and ϵ , we use this formula to determine N :

$$N = \left[\frac{Z_\alpha\sqrt{V\left(\theta\right)}}{\epsilon\overline{\theta}^*(P^*)} + \frac{D\left(\theta\right)}{\overline{\theta}^*(P^*)}\right]^2 \tag{20}$$

For given N and α , we use the following equation to solve for ϵ

$$\epsilon = \frac{Z_\alpha\sqrt{\frac{V(\theta)}{N}}}{\overline{\theta}^*(P^*) - \frac{D(\theta)}{\sqrt{N}}} \tag{21}$$

© 2012 NSP
Natural Sciences Publishing Cor.

Table 2: Display of standardized lower tolerance limit (γ) for given sample size N.

N	lower limit	standardized lower limit (in % 's)(γ)
100	3.06%	75%
200	3.35%	83%
300	3.48%	86%
400	3.55 %	88%
500	3.60 %	89%
600	3.64 %	90%
700	3.67 %	91%
800	3.70 %	91.4%
900	3.71%	92%
1000	3.73%	92.3%
1500	3.80 %	94%
2000	3.822%	94.5%
2100	3.827 %	94.7%
2200	3.832%	94.8%
2300	3.837%	94.9%
2350	3.839%	94.99%
2400	3.84%	95.04%

Table 4:Display of Standardized lower bound for given N ($\alpha = 0.05$ only)

2*N	Lower bound = $E(AI^w)(1-\epsilon)$	Standardized lower bound*
	Values of ϵ	γ
50	10%	66%
100	6.4%	76%
200	4.2%	83%
300	3.4%	86%
400	2.9%	88%
500	2.5%	89%

*Based on an alternative expression for the lower bound given by $\gamma\overline{\theta}^*(P^*)$.

5. Concluding Remarks

The concept of Apportionment index is relatively new. Still it transpires that eqivalent but different termenologies have been in use for quite some time. For example, Distribution Level (DL) is identical to the index AI. It is also known as Target Assortment Percentage(TAP) The well known Chi-square Statistic has been defined as usual. A variation of Chi-square, akin to DL, has also been introduced as Squared Distribution Level (SDL) and it is defined as

$$SDL = 100 * \left(1 - \frac{\sum_{i=1}^{k}(D_{di} - D_{oi})^2}{2}\right). \tag{22}$$

Malinen and Palander (2004) introduced what are called Penalty Segmented Distribution Level (PSDL) and Flexible Penalty Segmented Distribution Level (FPSDL), defined as

$$PSDL = 100 * \left(1 - \frac{\sum_{i=1}^{k} max\left(0, |D_{di} - D_{oi}| - T_i D_{di}\right)}{2}\right). \tag{23}$$

$$FPSDL = 100 * \left(1 - \sum_{i=1}^{k} \frac{(D_{di} - D_{oi})^2}{2F_i}\right). \tag{24}$$

We propose to study the statistical properties of these generalized versions of the Apportionment Index, with/without the price matrix under consideration.

We refer to Kirkkala et al (2000) and Malinen and Palander (2004) for studies related to these measures.

© 2012 NSP
Natural Sciences Publishing Cor.

6. Acknowledgement

The authors would like to thank (1) C R Rao AIMSCS for providing necessary facilities to Prof. B K Sinha's visit from ISI, Kolkata and (2) Head [BIRU], ISI, Kolkata for providing necessary facilities to Ms. J V Sukanya Divvela's visit from C R Rao AIMSCS. The work of the first author is funded under CMS DST Project Lr.No.SR/S4/MS:516/07, Dt.21-04-2008.

7. References

1. Agresti, A. (2002), Categorical Data Analysis (2nd ed.), New York, USA: Wiley.
2. Asikainen, A., Ala-Fossi, A., Visala, A. & Pulkkinen, P. (2005), "Metsteknologiasektorin isio ja tiekartta vuoteen 2020 [Vision and roadmap 2020 in forest technology]", Metlan tyraportteja 8, 91 pages. [In Finnish.]
3. Axelsson, S.-. (1998), "The mechanization of logging operations in Sweden and its effect on occupational safety and health", Journal of Forest Engineering, 9, 25–31.
4. Bergstrand, K.-G. (1989), "Frdelningsaptering med nroptimalmetoden – reviderad version [Bucking to order with a close-to-optimal method – revised version]", Forskningsstiftelsen Skogsarbeten, 1989-12-11, 11 p. [In Swedish.]
5. Drushka, K. & Konttinen, H. (1997), Tracks in the forest: the evolution of logging equipment, Helsinki, Finland: Timberjack Group Oy.
6. Gellerstedt, S. & Dahlin, B. (1999), "Cut-to-length: the next decade", International Journal of Forest Engineering, 10, 17–25.
7. Gellerstedt, S. (2002), "Operation of the single-grip harvester: motor-sensory and cognitive work", International Journal of Forest Engineering, 13, 35– 47.
8. Inman, H. F. & Bradley, E. L. (1989), "The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities", Communications in Statistics – Theory and Methods, 18, 3851–3874.
9. Jahn, J., Schmid, C. F. & Schrag, C. (1947), "The measurement of ecological segregation", American Sociological Review, 12, 293–303. Kiefer, J. (1975), "Construction and optimality of generalized Youden designs", in A survey of statistical design and linear models, ed. J. N. Srivastava, Amsterdam: North Holland Publishing Company, pp. 333–353.
10. Kirkkela, A., Sikanen, L., Harstela, P., Ruha, T & Tarnanen, T. (2000). Penalty segmented demand distribution for evaluating bucking outcome and calculating distribution level. *Sakkosegmentoitu tavoitejakauma apteeraustuloksen arvioinnissa ja jakauma-asteen laskennassa. Metstieteen aikakauskirja* 1/2000: 59-61.
11. Kivinen, V.-P. (2004), "A genetic algorithm approach to tree bucking optimization", Forest Science, 50, 696–710.
12. Kivinen (2006), "Design and testing of stand-specific bucking instructions for use on modern cut-to-length harvesters", Ph.D. thesis, Department of Forest Resource Management, University of Helsinki, Finland.
13. Kivinen, V.-P., Uusitalo, J. & Nummi, T. (2005), "Comparison of four measures designed for assessing the fit between the demand and output distributions of logs", Can. J. For. Res., 35, 693–702.
14. Koskela, L., Sinha, B.K., and Nummi, T. (2007). Some aspects of the sampling distribution of the Apportionment Index and related inference. *Silva Fennica* 41(1): 699–715.
15. Malinen, J. & Palander, T. (2004), Metrices for distribution similarity applied to the bucking to demand procedure, *International Journal of Forest Engineering* , 15, 33-40.
16. Marshall, H. D. (2005), "An investigation of factors affecting the optimal log output distribution from mechanical harvesting and processing systems", Ph.D. thesis, Oregon State University, Corvallis, OR.
17. Murphy, G. E. (2002), "Mechanization and value recovery: Worldwide experiences", in Proceedings of the Woodfor Africa 2002, Forest Engineering Conference: Forest Engineering Solutions for Achieving Sustainable Forest Resource Management – An International Perspective, Hilton College, Pietermanitzbury, South Africa, 2–3 July, 2002, eds. L. Kellog, B. Spong & P. Licth, Oregon State University, Department of Forest Engineering, Corvallis, Oregon, pp. 23–33.
18. Nummi, T., Sinha, B.K., and Koskela, L. (2005). Statistical Properties of the Apportionment Degree and Alternative Measures in Bucking Outcome. *Revista Investigacin Operacional*. Vol. 26, No. 3, 259-267.
19. Owende, P. M. O. (2004), "Wood Delivery", in Encyclopedia of Forest Sciences (Vol. 1), eds. J. Burley, J. Evans & J. A. Youngquist, Elsevier Ltd., Oxford, pp. 269–279.
20. Ponsse Oyj (2006), "Annual Report 2005", Ponsse Oyj, Kajaani, Finland.
21. Pulkki, R. (1997), "Cut-to-length, tree-length or full tree harvesting?", Central Woodlands, 1, n:o 3, 22–27, 37.
22. Sinha, B.K., Koskela, L., and Nummi, T. (2005a). On Some Statistical *Properties* of the Apportionment Index. *Revista Investigacin Operacional*, Vol. 26, No. 2, 169-179.
23. Sinha, B.K., Koskela, L., and Nummi, T. (2005b). On a Family of Apportionment Indices and Its Limiting Properties. *IAPQR Transactions*, Vol. 30, No. 2, 65-87.

© 2012 NSP
Natural Sciences Publishing Cor.

24.Weijo, A. (2000), ”Runkopankki Metshallituksen puunhankinnan ohjauksen apuvlineen [The stem bank as a control tool in the wood procurement of the Metshallitus corporation]”, Master’s thesis, Faculty of Forestry, University of Joensuu, Joensuu, Finland. [In Finnish.]

Appendix

Result A.1 : Suppose $X \sim N(\theta, \sigma^2)$.

Then

$$E(|X-\theta|) = \sigma\sqrt{\frac{2}{\pi}} and \; V(|X-\theta|) = \sigma^2(1-\frac{2}{\pi}). \tag{25}$$

The proof is omitted.

Result A.2: Suppose $X \sim N(\theta, \sigma^2)$

Then $V(X - |X-\theta|) = 2\sigma^2(1-\frac{1}{\pi})$.
Proof.

$$V(X-|X-\theta|) = V(X) + V(|X-\theta|) - 2Cov(X,|X-\theta|) = 2\sigma^2(1-\tfrac{1}{\pi}) - 2Cov(X,|X-\theta|).$$

Note that

$$\begin{aligned}
&Cov(X,|X-\theta|) = Cov(X-\theta,|X-\theta|)\\
&= E((X-\theta)\,|X-\theta|)\\
&= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(x-\theta)\,|X-\theta|\,e^{-\frac{1}{2}\left(\frac{x-\theta}{\sigma}\right)^2}dx\\
&= \sigma^2\int_{-\infty}^{\infty}u\,|u|\,\phi\,(u)\,du = 0.
\end{aligned}$$

since $u\,|u|$ is an odd function. Finally therefore,

$$V(X-|X-\theta|) = 2\sigma^2(1-\frac{1}{\pi}) - 2Cov(X,|X-\theta|) = 2\sigma^2(1-\frac{1}{\pi}). \tag{26}$$

Result A. 3. Suppose$(X,Y) \sim BVN\left(\theta_{1,}\theta_2,\sigma_1^2,\sigma_2^2,\rho\right)$.

Then $Cov\,(X,|Y-\theta_2|) = 0$.
Proof.

$$\begin{aligned}
Cov\,(X,|Y-\theta_2|) &= Cov\,(X-\theta_1,|Y-\theta_2|)\\
&= E[(X-\theta_1)\,|Y-\theta_2|]\\
&= E[E[(X-\theta_1)\mid(Y-\theta_2)]\,|Y-\theta_2|]\\
&= \rho\sigma_1/\sigma_2 E[(Y-\theta_2)\,|Y-\theta_2|] = 0.
\end{aligned} \tag{27}$$

by Result A.2 in (10).

Result A. 4.

In the following we assume$(X,Y) \sim BVN\left(\theta_{1,}\theta_2,\sigma_1^2,\sigma_2^2,\rho\right)$. Then

$$Cov\,(|X-\theta_1|\,,|Y-\theta_2|) = \sigma_1\sigma_2\varDelta\,(\rho)\,, \tag{28}$$

where $\varDelta\,(\rho) = \frac{2}{\pi}\left[\rho sin^{-1}\rho + \sqrt{(1-\rho^2)} - 1\right]$.
Proof.

© 2012 NSP
Natural Sciences Publishing Cor.

Sinha et al (2005) provided a proof of this result. Here we follow a different approach and give an alternative proof of this result.
We make the transformations $U = (X - \theta_1)/\sigma_1$, $V = (Y - \theta_2)/\sigma_2$ so that (U, V) follows $BVN(0, 0, 1, 1, \rho)$. Then, $Cov\left(|X - \theta_1|, |Y - \theta_2|\right) = \sigma_1\, \sigma_2 Cov\left(|U|, |V|\right) = \sigma_1 \sigma_2 \left[E\left(|U|\,|V|\right) - 2/\pi\right].$

We now proceed to evaluate an explicit expression for $E\left(|U|\,|V|\right)$. This we do by conditional argument, noting that conditional on $V = v$, U is normal with mean ρv and variance $(1 - \rho^2)$. Not to obscure any essential steps of reasoning, we proceed through the following steps.

Step 1

$$\begin{aligned}
\mathrm{E}(|U| \mid V) &= \textstyle\int_{-\infty}^{0} (-u)\, f(u \mid v) du + \int_{0}^{\infty} u f(u \mid v) du \\
&= \textstyle\int_{-\infty}^{\infty} u f(u \mid v) du - 2\int_{-\infty}^{0} u f(u \mid v) du \\
&= \textstyle\rho v - 2\int_{-\infty}^{0} u N(\rho v, \sqrt{(1-\rho^2)} du \\
&= \textstyle\rho v + 2\int_{0}^{\infty} u N(-\rho v, \sqrt{(1-\rho^2)} du. \qquad (28)
\end{aligned}$$

Step 2

The first term in the above will yield $E(V\,|V|) = 0$. As to the second term, we make usual variable transformation to $Z = (U + \rho v)/\sqrt{(1-\rho^2)}$. At this stage we set

$$c = \frac{\rho}{\sqrt{(1-\rho^2)}} \tag{29}$$

Then, ignoring the multiplier 2, the integral reduces to

$$\begin{aligned}
\int_{vc}^{\infty} (-\rho v + z\sqrt{(1-\rho^2)}\phi(z) dz &= (-\rho v)[1 - \Phi(vc)] + \sqrt{(1-\rho^2)} \int_{cv}^{\infty} z\phi\left(z\right) dz \\
&= -\rho v + \rho v \Phi(vc) + \sqrt{(1-\rho^2)}\phi(vc) \\
&= -T_1 + \rho T_2 + \sqrt{(1-\rho^2)} T_3. \qquad (30)
\end{aligned}$$

Note that $E(V\,|V|) = 0$ so that T_1 does not effectively contribute anything. Next, we evaluate $E\left(T_2\,|V|\right)$ and $E\left(T_3\,|V|\right)$ below.

Step 3.

$$E\left(T_2\,|V|\right) = E\left(V\Phi\left(Vc\right)|V|\right) \; = \int_{-\infty}^{\infty} v\Phi\left(vc\right)|v|\,\phi\left(v\right) dv \; = \int_{0}^{\infty} v^2\left(2\Phi(vc) - 1\right)\phi\left(v\right) dv \; = (2I^* - 1/2), say. \tag{31}$$

*E*valuation of I^*.

$$\begin{aligned}
I^* &= \int_0^{\infty} \left[v\left(\Phi\left(vc\right)\right)\right]\left[v\phi\left(v\right)\right] dv \\
&= \int_0^{\infty} \left(v\Phi\left(cv\right)\right)\left(v\left(\frac{1}{\sqrt{(2\pi)}} e^{-\frac{v^2}{2}}\right)\right) dv \\
&= \left[\left(v\Phi\left(cv\right)\right)\left[\frac{-1}{\sqrt{(2\pi)}} e^{-\frac{v^2}{2}}\right]\right]_0^{\infty} - \int_0^{\infty} \left(\Phi\left(cv\right) + cv\phi\left(cv\right)\right)\left(\frac{-1}{\sqrt{(2\pi)}} e^{-\frac{v^2}{2}}\right) dv \\
&= \int_0^{\infty} \Phi\left(cv\right)\left(\frac{1}{\sqrt{(2\pi)}} e^{-\frac{v^2}{2}}\right) dv + \frac{1}{\sqrt{(2\pi)}} \int_0^{\infty} \left(cv\right)\phi\left(cv\right) e^{-\frac{v^2}{2}} dv \\
&= \int_0^{\infty} \Phi\left(cv\right)\left(\frac{e^{-\frac{v^2}{2}}}{\sqrt{2\pi}}\right) dv + \frac{1}{\sqrt{(2\pi)}} \int_0^{\infty} c\left(\frac{e^{-\frac{v^2\left(c^2+1\right)}{2}}}{\sqrt{2\pi}}\right) d\left(\frac{v^2}{2}\right) \\
&= \int_0^{\infty} \Phi\left(cv\right)\phi\left(v\right) dv + \frac{c}{\left(c^2+1\right)2\pi} \\
&= \psi(c) + \frac{c}{\left(c^2+1\right)2\pi}, say. \qquad (32)
\end{aligned}$$

© 2012 NSP
Natural Sciences Publishing Cor.

$$I^* = \psi(c) + \frac{c}{(c^2+1)\,2\pi}. \tag{33}$$

We now deduce the expression for $\psi(c)$ below. Note that

$$\psi(c) = \int_0^\infty \Phi\,(cv)\,\phi\,(v)\,dv. \tag{34}$$

We treat this integral as a function of parameter c and, hence,

$$\psi^{'}(c) = \frac{\partial}{\partial c}\left[\int_0^\infty \Phi\,(cv)\,\phi\,(v)\,dv\right] \quad = \int_0^\infty \left(\frac{\partial}{\partial c}\Phi\,(cv)\right)\phi\,(v)\,dv \quad = \int_0^\infty v\phi\,(cv)\,\phi\,(v)\,dv \quad = \frac{1}{2\pi(1+c^2)}. \tag{35}$$

Hence, $\psi(c) = \frac{tan^{-1}c}{2\pi} + k$, k being the constant of integration.

Next, note that$\psi(0) = \int_0^\infty \Phi\,(0)\,\phi\,(v)\,dv = \frac{1}{4}$.

Therefore, $k + \frac{tan^{-1}0}{2\pi} = \frac{1}{4}$ i.e $k = \frac{1}{4}$.

Hence, finally

$$\psi(c) = \frac{tan^{-1}c}{2\pi} + \frac{1}{4} = \frac{sin^{-1}\rho}{2\pi} + \frac{1}{4}. \tag{36}$$

upon simplification.

Therefore,

$$E\,(T_2\,|V|) = 2\left[\left(\frac{sin^{-1}\rho}{2\pi} + \frac{1}{4}\right) + \frac{c}{(c^2+1)\,2\pi}\right] - \frac{1}{2}. \tag{37}$$

Step 4

Now we evaluate $E\,(T_3\,|V|)$ below.

$$E\,(T_3\,|V|) = E\,(\phi\,(cV)\,|V|) = 2\int_0^\infty v\phi\,(cv)\,\phi\,(v)\,dv = \frac{1}{\pi(1+c^2)}. \tag{38}$$

Step 5
Therefore, compiling all the results together, from expressions 14, 16, 23, 24
we obtain

$$\begin{aligned} E\,(|U|\,|V|) &= 2\rho E\,(T_2\,|V|) + 2\sqrt{(1-\rho^2)}E\,(T_3\,|V|) \\ &= 2\rho\left[2\left[\left(\frac{sin^{-1}\rho}{2\pi} + \frac{1}{4}\right) + \frac{c}{(c^2+1)\,2\pi}\right] - \frac{1}{2}\right] + \frac{2\sqrt{(1-\rho^2)}}{\pi(1+c^2)} \\ &= \frac{2}{\pi}\left(\rho sin^{-1}\rho + \sqrt{(1-\rho^2)}\right) \end{aligned} \tag{39}$$

upon simplification.

Therefore,

$$Cov\,(|U|\,|V|) = \frac{2}{\pi}\left[\rho sin^{-1}\rho + \sqrt{(1-\rho^2)} - 1\right]. \tag{40}$$

It is of interest to work out an expression for $Corr\,(|U|\,|V|)$ as well. This is indicated below.

© 2012 NSP
Natural Sciences Publishing Cor.

$$Corr\left(|U|\,|V|\right) = \frac{\frac{2}{\pi}\left[\rho sin^{-1}\rho + \sqrt{(1-\rho^2)} - 1\right]}{\left(1-\frac{2}{\pi}\right)} = \left(\frac{2}{\pi-2}\right)\left[\rho sin^{-1}\rho + \sqrt{(1-\rho^2)} - 1\right]. \tag{41}$$

Particular Cases

$Case1.\ \rho = 0 \Leftrightarrow c = 0.$
$Corr\left(|U|\,|V|\right) = \left(\frac{2}{\pi-2}\right)(1-1) = 0.$

$Case2.\ \rho = 1 \Leftrightarrow c = \infty.$
$Corr\left(|U|\,|V|\right) = \left(\frac{2}{\pi-2}\right)\left[sin^{-1}(1) - 1\right] = \left(\frac{2}{\pi-2}\right)\left[\frac{\pi-2}{2}\right] = 1.$

$Case3.\ \rho = -1 \Leftrightarrow c = -\infty.$
$Corr\left(|U|\,|V|\right) = \left(\frac{2}{\pi-2}\right)\left[\left\{(-1)\,sin^{-1}(-1)\right\} - 1\right] = \left(\frac{2}{\pi-2}\right)\left[(-1)\left(-\frac{\pi}{2}\right) - 1\right] = \left(\frac{2}{\pi-2}\right)\left[\frac{\pi-2}{2}\right] = 1$

© 2012 NSP
Natural Sciences Publishing Cor.