LAURA KOSKELA

Contributions to Statistical Aspects of Computerized Forest Harvesting

■

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty of Information Sciences of the University of Tampere,
for public discussion in the Paavo Koli Auditorium, Kanslerinrinne 1,
Tampere, on June 29th, 2007, at 12 o'clock.

UNIVERSITY OF TAMPERE

ACADEMIC DISSERTATION
University of Tampere
Department of Mathematics, Statistics and Philosophy
Finland

Distribution
Bookshop TAJU
P.O. Box 617
33014 University of Tampere
Finland

Tel. +358 3 3551 6055
Fax +358 3 3551 7685
taju@uta.fi
www.uta.fi/taju
http://granum.uta.fi

Cover design by
Juha Siro

Printed dissertation
Acta Universitatis Tamperensis 1237
ISBN 978-951-6982-4 (print)
ISSN 1455-1616

Electronic dissertation
Acta Electronica Universitatis Tamperensis 629
ISBN 978-951-44-6983-1 (pdf)
ISSN 1456-954X
http://acta.uta.fi

Tampereen Yliopistopaino Oy – Juvenes Print
Tampere 2007

Acknowledgements

I would like to thank my supervisor Dr. Tapio Nummi for introducing me to this interesting research area and offering me valuable ideas and insight into many problems encountered along the way to bringing this work to its final form. My thanks also for encouraging me to finalize my thesis.

I would also like to express my deepest gratitude to my assistant supervisor Professor Bikas K. Sinha, who has contributed immensely to my work during his regular visits to the University of Tampere. I warmly thank him for his willingess to share his wide expertise and a number of ideas throughout the preparation of the thesis. I am also grateful to him for arranging the opportunity to make an unforgettable and scientifically fruitful one month's visit to the Indian Statistical Institute (ISI), Kolkata, in July 2005.

I wish to thank Mr. V-P Kivinen for sharing his knowledge of forest harvesting and providing me with useful data. I also want to express my gratitude to him for reading the manuscript and providing valuable suggestions. I am indebted to the project leader Dr. Jori Uusitalo for sharing his expertise in forestry. My thanks also go to all the members of the project *Forest-level bucking optimization including transportation cost, product demands and stand characteristics* for successful co-operation. Ms. Simone Wenzel deserves my appreciation for her efforts in our joint work.

I would like to thank Professor Erkki P. Liski for giving me a helping hand whenever needed and for reading and commenting on the manuscript. I am also grateful to all my colleagues in my home department. A special thankyou to Ms. Anne Puustelli for always taking the time to listen to my thoughts.

My sincere thanks go to Mr. Robert MacGilleon for carrying out the greater part of the English language revision of this thesis.

For financial support I extend my warmest thanks to the Tampere Graduate School in Information Science and Engineering (TISE), the Niemi Foundation, the Scientific Foundation of the City of Tampere and the Academy of Finland. I am also grateful to the Department of Mathematics, Statistics and Philosophy in the University of Tampere for providing me with good research facilities.

And finally, most of all, I want to thank my friends and family for their efforts in keeping me normal and forcing me to take distance from my work. To my family I wish to express my deepest gratitude for the love and support they have always provided me. My parents and sisters – thank you for always being there for me! A special thankyou to my dear Jouni for his profound encouragement and support. In memoriam Mami († 4.6.2007).

Tampere, June 2007 *Laura Koskela*

Abstract

This thesis consists of six papers and a summary comprising statistical considerations of topics related to bucking optimization in cut-to-length forest harvesting. The topics addressed are: (1) the stem prediction problem in a harvesting situation and (2) measuring the fit between the demand and outcome distributions of logs. Since optimal tree bucking inevitably presumes accurate stem predictions, the choice of a proper stem prediction method is of crucial importance for the properties of all end products. Proper assessment of the bucking result has become relevant as the trend in the sawmill industry has been towards customer-oriented production of well-defined products.

The first article presents a cubic smoothing spline-based stem curve prediction and performs comparisons of this method with two other stem prediction techniques. In the second paper the use of a cubic smoothing spline is studied in the analysis of complete and balanced data. The basic idea was to replace the within-individual part of the Potthof and Roy GMANOVA model by cubic smoothing splines. It is shown how the mean splines can be estimated using a penalized log-likelihood function, and further, that the analysis can be greatly simplified under a certain special class of covariance structures. A rough testing of group profiles is also developed and illustrated.

The third paper studies the traditional χ^2-statistic in the context of measuring the bucking outcome and shows its relation to the Apportionment Index (AI) commonly used in harvesting in Scandinavia. The paper also presents price-weighted versions of both measures. The fourth paper examines the asymptotic sampling distribution of the AI by assuming a multinomial distribution for the bucking outcome. The paper provides approximate expressions for the first two moments of the measure and constructs the lower tolerance limit with a desired confidence level. In the first of the two remaining articles the definition of the AI and its price-weighted version are extended. The paper discusses the proper standardization of the measures and examines their limiting properties. The last article initiates a statistical analysis of the AI based on the joint distribution of random components in the outcome matrix. Dirichlet distribution is adopted to describe the joint distribution of the random components in the cases of two and three log categories. It is then proposed that the distribution parameters be chosen so that the AI is maximized in the averaged sense.

Key Words and Phrases: Apportionment Index, Cubic smoothing spline, Growth curve model, Measuring bucking outcome, Stem prediction.

Contents

Acknowledgements **3**

Abstract **4**

List of Original Publications **6**

1 Introduction **7**
1.1 Cut-to-length Harvesting 7
1.2 Bucking Optimization 8
1.2.1 Stem Level 9
1.2.2 Stand Level 10
1.3 Factors affecting the Bucking Outcome 10
1.4 Needs for Evaluating the Log Bucking Outcome 11

2 Stem Curve Prediction **12**
2.1 On Methods Proposed for Stem Curve Prediction for Harvesters 12
2.2 On Prediction Accuracy 15

3 Measuring the Bucking Outcome **17**
3.1 Target, Outcome and Price Matrix 17
3.2 Some Measures for Evaluating the Log Bucking Outcome . . . 18

Summaries of Original Publications **21**

List of Forestry Terms **26**

References **28**

A Supplementary to Paper III **32**
A.1 Approximating the Variance of the Price-weighted χ^2-statistic 32
A.2 Description of the Simulation Procedure 33
A.3 Justification for the Use of χ^2-distribution 33
A.4 Justification for the Use of the Weighted χ^2-distribution . . . 34

B Supplementary to Paper VI **36**

C Errata to the Original Papers **39**

List of Original Publications

I. Koskela, L., Nummi, T., Wenzel, S., and Kivinen, V.-P. (2006). On the Analysis of Cubic Smoothing Spline-Based Stem Curve Prediction for Forest Harvesters. *Canadian Journal of Forest Research*, Vol. 36, 2909-2920.

II. Nummi, T., and Koskela, L. (2006). Analysis of Growth Curve Data Using Cubic Smoothing Splines. Submitted to *Journal of Applied Statistics*.

III. Nummi, T., Sinha, B.K., and Koskela, L. (2005). Statistical Properties of the Apportionment Degree and Alternative Measures in Bucking Outcome. *Revista Investigación Operacional*. Vol. 26, No. 3, 259-267.

IV. Koskela, L., Sinha, B.K., and Nummi, T. (2007). Some Aspects of the Sampling Distribution of the Apportionment Index and Related Inference. Submitted to *Silva Fennica*.

V. Sinha, B.K., Koskela, L., and Nummi, T. (2005). On a Family of Apportionment Indices and Its Limiting Properties. *IAPQR Transactions*, Vol. 30, No. 2, 65-87.

VI. Sinha, B.K., Koskela, L., and Nummi, T. (2005). On Some Statistical Properties of the Apportionment Index. *Revista Investigación Operacional*, Vol. 26, No. 2, 169-179.

The papers are reproduced with the kind permission of the journals concerned. Each of the papers is summarized later in this thesis.

1 Introduction

The main part of this thesis consists of the six research papers listed in the previous chapter. The papers comprise statistical considerations of topics related to bucking optimization in cut-to-length (CTL) forest harvesting. In this chapter we provide a general introduction to bucking optimization in CTL harvesting and motivate the work of this thesis. Chapter 2 focuses on the prediction of stem profile, which plays an essential role in bucking optimization and serves as a topic for Paper I. The chapter also introduces a cubic smoothing spline function, which is used in the analysis of complete and balanced data in Paper II. Some methods used to evaluate the bucking outcome are discussed in Chapter 3. This chapter serves as an introduction and background especially for Papers III-VI. Section 3.2 contains a brief summary of the original papers. Some supplementary notes to Papers III and VI are presented in Appendix A and B, respectively. Appendix C contains errors perceived in some of the original papers. A list of forestry terms is provided for the reader not familiar with forestry vocabulary.

1.1 Cut-to-length 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

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).

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.

1.2 Bucking Optimization

Bucking tree stems optimally into various wood assortments and log lengths is one of the central and most challenging operations in the wood processing chain. Since a poor bucking outcome is hard or even impossible to compensate

later in the manufacturing process, the properties of all end products and thereby business profitability are crucially affected by the bucking process.

The basic principle in bucking optimization is to maximize the value of a single stem. However, since there are various market demands in terms of the amounts, types and characteristics of log products, maximizing the value of one single tree stem does not necessarily result in an optimal log output at stand level. It is therefore necessary to some extent to compromise on the principle of optimizing individual stems. In the following we first discuss bucking optimization at stem level, i.e. how the value of an individual stem is maximized. Then we broaden the perspective to stand level, where the aim is to determine an optimal bucking pattern not only for a single stem but for a large set of individual tree stems.

1.2.1 Stem Level

At stem level, the aim in bucking optimization is to assign to each harvested tree a bucking pattern which yields the highest total stem value (Kivinen, 2007). This principle is commonly called bucking-to-value.

Following the formulation of Liski & Nummi (1995), an admissible cutting pattern can be defined as a set of cutting points $0 = c_0 < c_1 < \ldots < c_K$ such that the length (l_k) and the small end diameter (d_k) of the kth log satisfy

$$l_k = c_k - c_{k-1} \in [l_{min}, l_{max}] \quad \text{and} \quad d_k \geq d_{min} > 0 \tag{1.1}$$

for $k = 1, 2, \ldots, K$, respectively, where c_0 is the cutting point at the butt of a tree, l_{min} is the minimum and l_{max} the maximum length of a log and d_{min} is the minimum acceptable log diameter. The marking for bucking problem (MBP) is defined by Näsberg (1985) as the problem of converting a single tree stem into smaller logs such that the total stem value according to a given price list is maximized for logs. The price list for a certain log product specifies how valuable or profitable it is to cut different length-diameter(-quality) combinations of particular log type and gives the price of a log as a function of both the length and the small end diameter (SED) of a log. The price of the whole stem is then the sum of single log prices. In general, the aim is to maximize a non-negative bounded utility function

$$H(c_0, c_1, \ldots, c_K) = \sum_{k=1}^{K} h(l_k, d_k) \tag{1.2}$$

under the constraints (1.1), where the function $h(l_k, d_k)$ can be taken as the price of the kth log from a given stem. Besides price, however, many other quantities (e.g. volume) can be used. Further, other constraints besides those in (1.1) are usually needed in practice.

The basic requirement for solving the marking for bucking problem optimally is that the whole stem profile be known and available to some sufficient level of accuracy during the bucking process. By stem profile (stem curve) is meant a function which describes how the stem tapering (diameter) changes with respect to the stem height. To know the stem curve before making cutting desicions, the whole stem could be first measured from the stump to the

top, then returning to the butt end to start bucking. However, in practice this kind of double processing is too slow and the stem could be damaged during the process. Modern harvesters are therefore equipped with a stem curve prediction system, which normally works in a stepwise manner. (1) After felling a tree, the harvester runs and measures it only for a length not exceeding the minimum log length, usually 3-4 m. (2) On the basis of the measurements of the tree currently processed and the profiles gathered from some number of previously cut stems, the harvester then predicts the profile of the unknown part of the stem and optimizes the cutting points. (3) At each suggested cutting point, the harvester usually checks whether the predicted stem diameter lies within the given tolerance from the diameter measured by the harvester at that point. If not, a new prediction and bucking optimization is performed, possibly changing the crosscutting point either backwards or forwards from its original place. Otherwise, the harvester cuts the log of the suggested length and, as more measured data on the stem are now available, updates the predicted profile for the remaining stem part and recalculates the further bucking points. Stem curve prediction is discussed in greater detail in Section 2.

1.2.2 Stand Level

The goal in stand-level bucking optimization is to assign a bucking policy which maximizes the aggregate production value from all stems being cut at a forest stand (Kivinen, 2007). Modern single-grip harvesters most frequently employ the bucking-to-demand (or bucking-to-order) principle, which incorporates both the log values and the desired log output distributions into the bucking optimization system. In bucking-to-demand optimization, a harvester, provided with the information on the value of each feasible length-diameter-quality combination of logs within each assortment (log product), selects the bucking pattern which maximizes the value of an individual stem (cf. stem level). However, besides selecting the bucking pattern with the highest overall value, the harvester also continuously monitors the difference between the mill's (or mills') demand log distribution and the actual output distribution. Two different implementations of bucking-to-demand optimization have been developed, namely the adaptive price list method and the close-to-optimal method. These methods are briefly introduced in Paper I. A more detailed description can be found e.g. in Kivinen (2007). Discussions of the optimization techniques and modeling approaches applied are provided, for example, in Marshall (2005) and Kivinen (2007).

1.3 Factors affecting the Bucking Outcome

Bucking optimization is a somewhat complex concept making it rather difficult to understand why optimization systems eventually produce different bucking outcomes in different circumstances. In an attempt to find answers to the above question, Uusitalo and Kivinen (2001) outline and discuss the most important factors affecting the bucking result with a modern tree buck-

ing optimizing system. These factors include e.g. measuring accuracy, stand composition, stem prediction accuracy, bucking algorithm, skill level of the harvester operator, demand distribution and the relationship between wood assortments.

As stated above, modern harvesters are equipped with a stem prediction system. Since inaccurate predictions commonly result in non-optimal bucking decisions, the prediction of stem profile is one of the most important parts of the bucking optimization system. The accuracy of the prediction method applied has a significant influence on the resulting bucking outcome both at the stem level and at the stand level. We devote Chapter 2 to the prediction of stem profile. For other factors we simply refer to the work of Uusitalo & Kivinen (2001).

1.4 Needs for Evaluating the Log Bucking Outcome

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, 2007). 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).

2 Stem Curve Prediction

2.1 On Methods Proposed for Stem Curve Prediction for Harvesters

Stem curve prediction in a harvesting situation when only a short part of the stem is known differs from the problem of modeling the whole stem curve. In stem curve prediction the main interest is in the unknown part of the tree. To make the distinction between these two tasks in the following discussion, we refer to the former by the expression "prediction" and to the latter by "modeling".

Before utilizing observed measurements for prediction purposes, computer programs often eliminate large intermittent errors in the diameter measurements, using e.g. filtering or smoothing techniques (e.g. Gellerstedt, 2002; Lukkarinen & Marjomaa, 1997). Let y_i denote the observed and smoothed stem diameter at point x_i, $i = 1, 2, \ldots, m$, where x_i is the distance of the ith measurement from the butt and m is the total number of measurements. We assume that

$$y_i = d(x_i) + \epsilon_i, \tag{2.1}$$

where the stem curve $d(x_i) = \mathrm{E}(y_i)$ is a smooth decreasing function in stem height x_i and ϵ_i is random error. The prediction problem is to determine the stem curve measurements at the forthcoming stem points $x_{m+1}, x_{m+2}, \ldots, x_n$. Bucking optimization is then based on the observed stem curve measurements $y_1, y_2, \ldots, y_m$ and on the predictions $\hat{y}_{m+1}, \hat{y}_{m+2}, \ldots, \hat{y}_n$.

The advanced prediction methods are based on mathematical models. Although stand density, site type, climate, genetic factors etc. are known to affect the form of the stem (Laasasenaho, 1982, p. 18), all such variables cannot usually be included in stem curve models, since in practice they are either difficult or impossible to measure.

The parameters of the model are commonly estimated by a set of previously harvested trees (data window) and possibly also by the measured part of the stem being processed. As a new tree has been harvested, the data window is updated by removing the oldest stem and adding the newly harvested tree. The size of the data window is kept small to adapt to possible changes in the stem population (see e.g. Liski & Nummi, 1995). Different prediction methods have been developed for harvesters, especially in Scandinavia. Since harvester manufactures have been unwilling to publish very detailed information on the methods actually utilized in harvesters, the number of well-documented prediction methods is small.

One of the simplest stem prediction methods based on a mathematical model utilizes a linear curve and at a minimum two diameter measurements taken at different heights on the known part of the stem (see Lukkarinen & Marjomaa, 1997). The measurements are usually chosen such that the unfavorable effect of the irregular butt section and butt swelling could be avoided. Assuming that the tapering of the rest of the stem can be adequately described by the recorded measurements, the unknown part is predicted by drawing a straight line through the two chosen points. Some more advanced versions of this method take into account the height at which prediction is calculated and then make different corrections. The method is simple and fast and requires only a small amount of computational power. Its main disadvantage is sensitivity to measurement errors and irregularities in stem shape. It also often fails to describe the butt end and top of the tree adequately.

The most advanced stem profile prediction methods are based on relative stem shape theory or the so-called mixed model techniques. According to the relative stem shape theory the taper curve in different-sized trees of the same species is of the same shape and the absolute variation caused by differences in tree size is eliminated by modeling relative diameters at relative heights along the stem (e.g. Laasasenaho, 1982; Kozak, 1988; Newnham, 1992). Laasasenaho (1982) suggested a polynomial model, in which the dependent variable is the ratio of the stem diameter to the diameter at 20% of tree height and the independent variable is the relative height. The powers used in the polynomial model are in accordance with the Fibonacci series and the model can be written as

$$\frac{d_l}{d_{,2h}} = b_1 x + b_2 x^2 + b_3 x^3 + b_4 x^5 + b_5 x^8 + b_6 x^{13} + b_7 x^{21} + b_8 x^{34}, \tag{2.2}$$

where $d_{,2h}$ is the basic diameter at 20% height ($x = 0.8$), d_l is the diameter at a height of l from the ground, $x = 1 - \frac{l}{h}$ or the relative distance from the top (0 corresponds to the top and 1 to the butt of the tree) and b_i ($i = 1, \ldots, 8$) are the model parameters. It has been noted that models with only the first 5 or 6 terms seem to suffice in a real harvesting situation. Laasasenaho's (1982) model has been utilized in the stem prediction methods of some Finnish harvester manufacturers (Lukkarinen & Marjomaa, 1997).

Since the real stem height is not available at the time of bucking and individual variation in the form of stems is not usually perceived in the models, relative stem shape theory-based models may be difficult to adapt to stem prediction in harvesting. It is of course possible to use modifications whereby the unknown values of the variables in stem curve equations are predicted. However, the parametric form of such equations depends crucially on the accuracy of the predicted variables. Poor prediction for stem height, for example, may ruin the form of the complete stem curve. It is further known that parameters in such models may not be unbiasedly estimated using standard estimation procedures when the variables in the stem equation are measured with error. For stem curve prediction when the stem height is measured with error we refer to Nummi & Möttönen (2004b).

Lappi (1986) used linear mixed models for modeling stem curves with the dimensions of a tree stem defined by a polar coordinate system. Liski &

Nummi (1995) studied polynomial mixed models for repeated measurements for stem prediction based on the real stem data. The new feature in the latter was the incorporation of the individual form variation of stems. The authors suggested the use of the second-degree polynomial with two random effects. The model for a stem at height x_i can be written as

$$y_i = (\beta_0 + b_0) + (\beta_1 + b_1)x_i + \beta_2 x_i^2 + \epsilon_i, \tag{2.3}$$

where β_0, β_1 and β_2 are common mean curve parameters, b_0 and b_1 are random effects associated with the individual stem to be predicted and ϵ_i is an error term. The model assumes that parameters b_0 and b_1 and random errors are independent and independently and identically normally distributed. According to Lukkarinen & Marjomaa (1997) a modification of this method with a third-degree polynomial has been applied e.g. in Ponsse Opti systems.

One slight drawback in the approach of Liski and Nummi, however, is that although most stem curves can be well predicted by low-degree random coefficient polynomial models (unknown part), in certain cases the fit may be rather poor. This may be the case where the butt of the tree is large and irregular. A possible extension is to use non-linear mixed models (e.g. Eerikäinen, 2001; Garber & Maguire, 2003). However, the use of these for prediction in a harvesting situation is not, according to our knowledge, well established.

The approaches of Laasasenaho (1982) and Liski and Nummi (1995), for example, constrain the curve estimates to certain pre-specified parametric forms, i.e. polynomials. A relatively straightforward extension of parametric regression modeling is the use of spline functions. Suppose now that we have K distinct points on some interval $[a, b]$ and refer to these points as knots. For example, knots could be some K stem points $x_1, x_2, \ldots, x_K$ on an inteval $[a, b]$ satisfying $a < x_1 < x_2 < \ldots < x_K < b$. The spline of order p with the given knots $x_1, x_2, \ldots, x_K$ can be written in the form

$$d(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \ldots + \beta_p x^p + \sum_{k=1}^{K} u_k (x - x_k)_+^p,$$

where

$$(x - x_k)_+ = \begin{cases} 0, & x \leq x_k \\ x - x_k, & x > x_k \end{cases}$$

and $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \ldots, \beta_p)'$ and $\mathbf{u} = (u_1, \ldots, u_k)'$ denote vectors of coefficients and $1, x, x^2, x^3, (x - x_1)_+^3 \ldots, (x - x_k)_+^p$ are called basis functions. The equation describes a sequence of p degree polynomials tied together at the knots to form a continuous function. Quadratic ($p = 2$) and cubic ($p = 3$) splines are the most commonly used splines. A natural cubic spline is obtained by assuming that the function is linear beyond the boundary knots x_1 and x_K. Assuming that the error terms are independent with common mean zero and variance σ_ϵ^2, the coefficients $\boldsymbol{\beta}$ and $\mathbf{u}$ can be estimated using the standard least squares method.

In some cases the use of the standard least squares procedure may result in a very rough curve estimate. To control this roughness, one may use the

so-called penalized sum of squares procedure, where the curve estimate is determined not only by its goodness-of-fit to the data as quantified by the least squares function but also by its roughness. The roughness of a twice-differentiable curve $d(\cdot)$ at interval $[a, b]$ can be measured, for example, by calculating its integrated squared second derivative $\int_a^b \{d''(t)\}^2 dt$. Suppose now that $y_1, y_2, \ldots, y_K$ are the observed values at the knots. A (natural) cubic smoothing spline $d(\cdot)$ is a smooth and continuously twice-differentiable curve which (for fixed α) minimizes the penalized sum of squares,

$$\sum_{i=1}^{K} \{y_i - d(x_i)\}^2 + \alpha \int_a^b \{d''(x)\}^2 dx,$$

where α is a positive smoothing parameter which controls the smoothness of the curve (see e.g. Green & Silverman, 1994). For large values of α, the curve estimate will display very little curvature and in the limiting case as α tends to infinity, the spline curve will approach the linear regression fit. For relatively small values of α, the curve estimate will track the data closely and in the limiting case as the parameter value tends to zero, the spline curve will approach the natural cubic smoothing spline which interpolates the data points (c.f. interpolating cubic spline). Thus, by controlling the smoothing parameter value we may adjust the fit smoothly from linear regression to natural cubic spline.

The idea of using splines in stem curve modeling is not particularly new. Stem curve models have been built using an interpolating cubic spline on the basis of several diameter measurements (e.g. Lahtinen & Laasasenaho, 1979; Goulding, 1979; Figueiredo-Filho et al., 1996) and e.g. so-called monotony-preserving taper curves have been constructed using a quadratic spline (Lahtinen, 1988). Cubic smoothing splines were employed to portray the stem curve for example in a study by Liu (1980). Recently, spline-based techniques have yielded promising results as applied to prediction of stem curve in a harvesting situation. The idea of applying cubic smoothing splines in this context was first introduced by Möttönen & Nummi (2002). The method utilized branch limits, which are difficult or impossible to measure in a real harvesting situation. One advantage of the method is that it does not assume any special functional form for the whole stem curve, although some models are needed to predict the branch limits. The model was subsequently modified by Nummi & Möttönen (2004a) using multivariate regression models with smoothing splines.

2.2 On Prediction Accuracy

One way to assess the performance of a prediction method is to compare the total stem values of a prediction-based bucking process with the optimal values obtained by utilizing the complete stem curves. Based on this idea, Näsberg (1985) found that loss in value due to incomplete stem information (i.e. a part of the stem was measured and the remaining part was predicted) was less than 2%. These good results were partially explained by the fact that

the quality limits and defect positions of the processed trees were assumed to be known precisely in the test. In a study by Liski & Nummi (1995), using the mixed-effects model based prediction method, the minimum percentage loss in value was found to be about 5%. The accuracy of a prediction method based on Laasasenaho's (1982) model was studied for spruce by Vuorenpää et al. (1997). Using the Apportionment Index (see Equation 3.1) as the criterion for bucking performance, a less than five percentage unit smaller index value was obtained for prediction-based bucking than when the cutting was based on complete stem measurements.

Lukkarinen & Marjomaa (1997) studied the prediction accuracy of the Ponsse harvester on both spruce and pine stems, obtaining better results for pine. The deviation between the predicted and measured SED of logs was on average approximately +3 mm for spruce and -9 mm for pine. The standard deviation for spruce and pine was 13 mm and 21 mm, respectively. The average error obtained when predicting length was on average +11 cm for spruce and -28 cm for pine. The respective standard deviations were 140 cm and 160 cm. The study by Lukkarinen et al. also showed that somewhat better predictions are obtained as more is known of the stem under process. Similar conclusions were drawn by Liski & Nummi (1995) and Marshall (2005).

3 Measuring the Bucking Outcome

3.1 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

$$\mathbf{T} = (t_{ij}) = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ t_{m1} & t_{m2} & \cdots & t_{mn} \end{bmatrix}$$

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 ith diameter class and jth length class, $i = 1, \ldots, m$ and $j = 1, \ldots, n$. 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 $d_i \leq d < d_{i+1}$ and $l_j \leq l < l_{j+1}$. Correspondingly, $m \times n$ matrix

$$\mathbf{O} = (o_{ij}) = \begin{bmatrix} o_{11} & o_{12} & \cdots & o_{1n} \\ o_{21} & o_{22} & \cdots & o_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ o_{m1} & o_{m2} & \cdots & o_{mn} \end{bmatrix}$$

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

$$\mathbf{P} = (p^*_{ij}) = \begin{bmatrix} p^*_{11} & p^*_{12} & \cdots & p^*_{1n} \\ p^*_{21} & p^*_{22} & \cdots & p^*_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p^*_{m1} & p^*_{m2} & \cdots & p^*_{mn} \end{bmatrix},$$

where $p^*_{ij} = \frac{p_{ij}}{\sum_{k=1}^{m} \sum_{l=1}^{n} p_{kl}}$ is the relative price of the ith diameter and jth length combination of logs and p_{ij} is the respective absolute price.

3.2 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[1] in forestry by Bergstrand in the mid-1980s (e.g. Bergstrand, 1989). For a fixed quality class the AI is defined as

$$AI = 1 - 0.5 \times \sum_{i=1}^{m} \sum_{j=1}^{n} |o_{ij}^* - t_{ij}^*|, \tag{3.1}$$

where $o_{ij}^* = \frac{o_{ij}}{\sum_{k=1}^{m} \sum_{l=1}^{n} o_{kl}}$ and $t_{ij}^* = \frac{t_{ij}}{\sum_{k=1}^{m} \sum_{l=1}^{n} t_{kl}}$ 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

$$AI = \sum_{i=1}^{m} \sum_{j=1}^{n} \min(o_{ij}^*, t_{ij}^*). \tag{3.2}$$

The maximum value of the AI is 1 (100%), which indicates a perfect match between the distributions. The minimum value of the index is $\min(t_{11}^*, t_{12}^*, \ldots, t_{mn}^*)$, 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. $c_{ij} = o_{ij} - t_{ij}$, the AI can also be expressed for equal matrix totals as

$$AI = \frac{N - \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} I(c_{ij} > 0)}{N}, \tag{3.3}$$

where $N = \sum_{i=1}^{m} \sum_{j=1}^{n} o_{ij} = \sum_{i=1}^{m} \sum_{j=1}^{n} t_{ij}$ 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.

[1]The history of some measures closely related to the AI (e.g. Dissimilarity Index discussed later in this section) strongly suggest that the AI was not developed by the Swedish mathematician Bergstrand in the mid-1980s, as claimed in many forestry papers, but had appeared earlier in different contexts such as sociology.

(2005) and Nummi et al. (2005) [Paper III]. The price-weighted Apportionment Index utilizes the price matrix and is defined as

$$AI_p = \sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij}^* \min(o_{ij}^*, t_{ij}^*). \tag{3.4}$$

The AI_p 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. However, in view of the scope of this thesis these measures are not discussed here.

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 χ^2-test, which was applied in the forestry context e.g. in Malinen & Palander (2004), Kivinen et al. (2005) and Nummi et al. (2005) [Paper III]. Using the same notations as above, the test statistic can be defined as

$$\chi^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{(o_{ij} - t_{ij})^2}{t_{ij}}.$$

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 χ^2-statistic by using the contingency coefficient C defined as

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}},$$

where N is the total number of logs harvested. Substracting 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) [Paper III], 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 $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ are density functions defined on the n-dimensional Euclidian space R_n, then the OVL can be defined as

$$OVL = \int_{R_n} \min[f_1(\mathbf{x}), f_2(\mathbf{x})] d\mathbf{x}.$$

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 χ^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 thesis, we may content ourselves with a reference to this particular study.

Summaries of Original Publications

I. In the article in question we briefly introduce the bucking process and then mathematically formulate the stem prediction problem. We also briefly outline cubic smoothing splines and present a cubic smoothing spline-based stem curve prediction method applicable in a real harvesting situation. The method is based on the idea first introduced in Möttönen & Nummi (2002) and subsequently modified by the authors in Nummi & Möttönen (2004a).

   The performance of the cubic smoothing spline-based method is compared to the linear mixed model approach of Liski & Nummi (1995) and the stem curve prediction method based on Kozak's taper equation (Kozak, 1988). For comparisons we use a study material consisting of five sets of Scots pine and Norway spruce stem profiles collected by two different harvester models in five different final felling stands in southern Finland. The performance of the methods is assessed by studying the prediction errors by means of Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). The results of comparisons show that the spline-based approach outperforms the other two methods. For example, the MAPE values of the spline-based method vary from 2.1% to 4%, while for the linear mixed model-based approach and the method based on Kozak's taper equation the values vary from 4.6% to 7.0% and from 2.7% to 4.7%, respectively.

II. A common approach used to model longitudinal data, i.e. data where individuals are measured according to some ordered variable, is based on the linear mixed models for repeated measures. Although this model provides an eminently flexible approach to modeling a wide range of mean and covariance structures, it is forced into a rigidly defined class of mathematical formulas which may not be well supported by the data within the whole sequence of observations. A cubic smoothing spline provides a non-parametric alternative to modeling such data. It can be shown that under normality assumption the solution of the penalized log-likelihood equation is the cubic smoothing spline, and this solution can be further expressed as a solution of the linear mixed model (see e.g. Green & Silverman, 1994 and Verbyla et al., 1999). As the first result, we show that the simple unweighted estimator can be used instead of the weighted estimator when the covariance of errors belongs to a certain special class of covariance structures, which assume particular importance when splines are used to analyse a group of individuals.

According to our knowledge this result is new in the smoothing spline context.

The main part of the paper is devoted to showing how cubic smoothing splines can be easily used in the analysis of complete and balanced data. The basic idea is to replace the within-individual part of the Potthof and Roy (1964) model GMANOVA (Generalized Multivariate Analysis of Variance) by cubic smoothing splines. It is then shown how the mean splines can be estimated using a penalized log-likelihood function. It is further shown that the analysis can be greatly simplified under a certain special class of covariance structures discussed earlier in the paper. The connection to mixed models is used in developing the rough testing of group profiles and numerical examples are presented to illustrate the techniques proposed.

III. Testing statistically that the distribution of the population from which the data is drawn agrees with a posited distribution is constantly encountered in many areas of research. The most commonly used test for examining the goodness-of-fit of grouped data is the frequency χ^2-test. In this paper we study the use of the χ^2-statistic in the context of measuring the goodness of the bucking outcome and show its relation to the Apportionment Index traditionally used in practice in Scandinavia. Since the Apportionment Index is often criticized as not taking account of the price deviation between different log categories, i.e. it gives the same weight for all log categories, we also introduce price-weighted versions of both measures.

Applying the large sample properties of the frequency χ^2-distribution we justify the use of the weighted χ^2-distribution as an approximation to the distribution of the price-weighted χ^2-statistic. A simulation study is conducted to illustrate the behaviour of the measures when there is a shortfall of a fixed proportion of logs in the outcome matrix with respect to the target. The simulation shows that even large proportions of missing logs ($\approx 40\%$) may give relatively high values of both the traditional and the non-weighted AIs, while the χ^2-statistic and its price-weighted versions will reject the hypothesis of agreement between the distributions as 20% or more of the logs are missing. This indicates better statistical performance of the χ^2-statistic and its price-weighted version. Since in practice it may not be possible to attain the demanded values and/or the required total number of logs exactly and virtually small deviations from the target values may not be of much interest, the behaviour of both AI measures may be more desirable for the practical applications. Note that the AI and its price-weighted version only compare the relative values of the observed and demanded distributions and therefore even large departures in the absolute values may not be noticed.

Appendix A contains some notes supplementary to Paper III. In Appendix A.1 we justify the statements of the approximation of the variance of the $\chi^2(p^*)$ given in the original paper. In Appendix A.2 we

clarify the arrangements of the simulation study conducted in Section 3 in the paper. Appendix A.3 and A.4 justify the use of the non-weighted and weighted χ^2-distributions for the standard and price-weighted χ^2-measures, respectively. Errata to the original paper are provided in Appendix C.

IV. In this paper we examine the asymptotic sampling distribution of the Apportionment Index by assuming a multinomial distribution for the bucking outcome. Such an assumption is natural in the context of a frequency distribution. Under the multinomial assumption and using large-sample normal approximations, we derive the approximate expressions for the first and second moment of the AI and construct the lower tolerance limit with a desired confidence level. A simulation study is then carried out to evaluate the accuracy of the approximations. The determination of the number of logs needed to attain high apportionment with a given accuracy is also studied. The effects of the number of logs harvested as well as of the form and the size of the given target matrix are discussed in the paper.

The formulas of the first two moments clearly show that under the multinomial assumption both the expected value of the AI and the variance depend on the number of logs harvested (N). Since N as well as the form and the size of the given target matrix seem to affect the index value (as discussed in Section 5 in the paper), some justification for a tolerable index value is needed. Here we suggest AI values higher than the mean $E(AI)$ to indicate a satisfactory level of agreement. If the AI falls below the lower tolerance limit, we would consider the outcome not satisfactory from the point of view of agreement with the given target.

As the AI simply gives the proportion of the "correctly" located logs in the outcome matrix with respect to the given target, we propose to examine the log categories in the outcome matrix which indicate "upload" and "download" with respect to the target. This is justified when we note that the computation of the AI depends only on the total amount of upload (or download) and not on its specific frequency distribution. However, we have not attempted to answer the question of a desirable distribution for the upload and download. This calls for a thorough study and close interaction with forestry personnel, as not only the price matrix but also e.g. the dimensions of the logs indicating upload and download may have some crucial impact on the overall agreement.

V. Here we extend the definition of the AI and its price-weighted version and discuss proper standardizations of these measures. Applying some aspects of moments and the Liaponouv inequality we then examine the limiting properties of the measures and provide some examples to illustrate the behaviour of the non-weighted and price-weighted generalizations.

The idea of introducing a Family of Apportionment Indices derives from the work done in the field of optimal design of experiments initiated by Kiefer (1975), who proposed a family of optimality criterion which includes the well-known A-, D- and E-optimality criteria as special cases. In this paper we show that the generalized Apportionment Index has its connection to e.g. harmonic, geometric and arithmetic means. One possible application of this kind of study might be to examine which of two or more competing outcome matrices is more robust with respect to the aspect of closeness to a given target matrix.

VI. The paper initiates a statistical analysis of the Apportionment Index based on the joint distribution of random component outputs in the outcome matrix. Dirichlet distribution is adopted to describe the joint distribution of the random components. This is justified by noting that all terms in the outcome matrix belong to the interval [0,1] and they add up to one. Our purpose is then to choose the parameters of the distribution so that the maximum apportionment is achieved. Here we propose to maximize the AI in the averaged sense, i.e. we aim at maximizing its expected value. Another approach would be, for example, to attain heavy right-tail distribution for the AI such that the index value would tend to be probabilistically large. However, we will not pursue the latter approach in this thesis.

We first study the case of only two log categories in which case maximizing the AI amounts to minimizing the mean deviation for any one of the two log categories (as the other is determined automatically). Since the mean deviation is least when it is taken about the median of the distribution, our goal is to identify the parameter values for which the median is the known target value. However, some condition is needed to find a unique solution to the problem. Here we stipulate a condition on the variance of the random outputs by assuming that some bound is desirable on the accuracy of the bucking outcome. Demonstrations of the specification of the parameters of the beta distribution (a special case of Dirichlet distribution) are given when the target and an upper bound to the variance of the component(s) are pre-specified.

We also extend the analysis to the case of three log categories aiming at specifying the parameter values of the underlying Dirichlet distribution by maximizing the expected AI such that the mean deviations between the random components and the respective target values are simultaneously minimized. This calls for minimizing the mean deviation such that the target value is taken as the median of the marginal distribution of the respective random outcome. However, as argued in the paper, simultaneous minimization of all three terms is not possible. A method is then proposed to tackle the problem of specifying the parameter values as the target matrix is given and an upper bound to the largest of the variances of the random outputs is specified. By doing so, we also have control over the variablility of other two random outputs.

As shown in the paper, the specification of the parameter values by

maximizing the AI in the averaged sense under the proposed Dirichlet distribution for the random outputs is a complex task and it is not amenable to an analytical solution even in the case of three locations. Generalization of the method initiated in this paper to more than three log categories clearly calls for a thorough investigation with powerful computational tools. However, if only some parts of the target matrix are of special interest in harvesting planning (e.g. large and/or small target values), a study of the appropriate submatrices may well suffice and the method provided in this paper could hence come into consideration.

Appendix B contains some notes supplementary to Paper VI. Since the appropriateness of the technique proposed to tackle the problem of specifying the parameter values in the case of three locations is not studied in the paper, we now take up some computations on the suitability of the proposed technique. Errata to the original paper are provided in Appendix C.

List of Forestry Terms

The literature citations referenced for each term follow the definition (in brackets). The following numbering is used for the citations: [1] Stokes et al. (1989), [2] Dykstra & Heinrich (1995) and [3] Megalos & Kea (2003). These references can be used as sources for terms not found in this glossary.

Buck	To saw a felled tree into short cuts. [1]
Bucking	The act or process of transversely cutting the stem or branches of a felled tree into logs. [2]
Butt	Base of a tree. Large end of a log. [1]
Crosscutting	See bucking. [2]
Cutting	Process of felling trees. [1]
Delimbing	Removing branches from trees. [1]
Felling	The act or process of severing a standing tree. Compare cutting. [2]
Forwarder	Self-propelled or mobile machine, usually self-loading, designed to transport trees or parts of trees by carrying them completely off the ground. [1]
Forwarding	Transporting trees or parts of trees by carrying them completely off the ground rather than by pulling or dragging them along the ground. [1]
Haul	Convey wood from a loading point to an unloading point. [1]
Harvester	A machine which fells trees, delimbs them and crosscuts them into logs. [2]
Harvesting	The aggregation of all operations, including pre-harvest planning and postharvest assessment, related to the felling of trees and the extraction of their stems or other usable parts from the forest for subsequent processing into industrial products. Also called timber harvesting. [2]
Log	Length of tree suitable for processing into lumber, veneer, or other wood products. [1]
Logging	The act or process of felling and extracting timber from forests, especially in the form of logs. [2]

Stand	An easily defined area of forest which is relatively uniform in species composition or age and can be managed as a single unit. [3]
Stem	Main body of a tree from which branches grow. Used loosely to refer to trees. [1]
Stump	The woody base of a tree remaining in the ground after felling. [2]
Timber	General term applied to forests and their products. [1]
Topping	Cutting off the top of a tree at a predetermined, minimum diameter. [1]
Yard	Place where logs are accumulated. [1]

References

Agresti, A. (2002), *Categorical Data Analysis* (2nd ed.), New York, USA: Wiley.

Asikainen, A., Ala-Fossi, A., Visala, A. & Pulkkinen, P. (2005), "Metsäteknologiasektorin visio ja tiekartta vuoteen 2020 [Vision and roadmap 2020 in forest technology]", Metlan työraportteja 8, 91 pages. [In Finnish.]

Axelsson, S.-Å. (1998), "The mechanization of logging operations in Sweden and its effect on occupational safety and health", *Journal of Forest Engineering*, 9, 25–31.

Bergstrand, K.-G. (1989), "Fördelningsaptering med näroptimalmetoden – reviderad version [Bucking to order with a close-to-optimal method – revised version]", Forskningsstiftelsen Skogsarbeten, 1989-12-11, 11 p. [In Swedish.]

Bilodeau, M. & Brenner, D. (1999), *Theory of multivariate statistics*, New York: Springer-Verlag.

Drushka, K. & Konttinen, H. (1997), *Tracks in the forest: the evolution of logging equipment*, Helsinki, Finland: Timberjack Group Oy.

Dykstra, D. P. & Heinrich, R. (1995), "FAO model code of forest harvesting practice", Food and Agricultural Organization of the United Nations, Rome, Italy, 117 p. Available from Internet: `http://www.fao.org/docrep/v6530e/v6530e12.htm` [accessed 9 December 2006].

Eerikäinen, K. (2001), "Stem volume models with random coefficients for Pinus kesiya in Tanzania, Zambia, and Zimbabwe", *Can. J. For. Res.*, 31, 879–888.

Figueiredo-Filho, A., Borders, B. E. & Hitch, K. L. (1996), "Number of diameters required to represent stem profiles using interpolated cubic splines", *Can. J. For. Res.*, 26, 1113–1121.

Garber, S. M. & Maguire, D. A. (2003), "Modeling stem taper of three central Oregon species using nonlinear mixed effects models and autoregressive error structures", *For. Ecol. Manage.*, 179, 507–522.

Gellerstedt, S. (2002), "Operation of the single-grip harvester: motor-sensory and cognitive work", *International Journal of Forest Engineering*, 13, 35–47.

Gellerstedt, S. & Dahlin, B. (1999), "Cut-to-length: the next decade", *International Journal of Forest Engineering*, 10, 17–25.

Goulding, C. J. (1979), "Cubic spline curves and calculation of volume of sectionally measured trees", *N.Z. J. For. Sci.*, 9, 89–99.

Green, P. J. & Silverman, B. W. (1994), *Nonparametric regression and generalized linear models*, London: Chapman & Hall.

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.

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.

Kirkkala, A., Sikanen, L., Harstela, P., Ruha, T. & Tarnanen, T. (2000), "Sakkosegmentoitu tavoitejakauma apteeraustuloksen arvioinnissa ja jakauma-asteen laskennassa [Assessing the goodness of the bucking outcome by penalty-segmented log demand distribution]", *Metsätieteen aikakauskirja 1/2000*, pp. 59–61. [In Finnish.]

Kivinen, V.-P. (2004), "A genetic algorithm approach to tree bucking optimization", *Forest Science*, 50, 696–710.

——— (2007), "Design and testing of stand-specific bucking instructions for use on modern cut-to-length harvesters", Academic dissertation, Dissertationes Forestales 37, Department of Forest Resource Management, University of Helsinki, Finland.

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.

Kozak, A. (1988), "A variable-exponent taper equation", *Can. J. For. Res.*, 18, 1363–1368.

Laasasenaho, J. (1982), "Taper curve and volume functions for pine, spruce and birch", *Commun. Inst. For. Fenn.*, 108, 1–74.

Lahtinen, A. (1988), "On the construction of monotony preserving taper curves", *Acta For. Fenn.*, 203, 1–34.

Lahtinen, A. & Laasasenaho, J. (1979), "On the construction of taper curves by using spline functions", *Commun. Inst. For. Fenn.*, 95, 1–63.

Lappi, J. (1986), "Mixed linear models for analyzing and predicting stem form variation of Scots pine", *Commun. Inst. For. Fenn.*, 134, 1–69.

Liski, E. P. & Nummi, T. (1995), "Prediction of tree stems to improve efficiency in automatized harvesting of forests", *Scandinavian Journal of Statistics*, 22, 255–269.

Liu, C. J. (1980), "Log volume estimation with spline approximation", *For. Sci.*, 26, 361–369.

Lukkarinen, E. & Marjomaa, J. (1997), "Rungon kapenemisen ennustaminen hakkuukoneen mittalaitteella – Ennustusmenetelmien vertailu", Metsätehon raportti 35, Metsäteho Oy, Helsinki, Finland, 29 p. [In Finnish.]

Malinen, J. & Palander, T. (2004), "Metrics for distribution similarity applied to the bucking to demand procedure", *International Journal of Forest Engineering*, 15, 33–40.

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.

Megalos, M. A. & Kea, J. B. (2003), "Understanding forestry terms: a glossary for private landowners", Woodland Owner Notes, North Carolina Cooperative Extension Service, North Carolina State University. Available from Internet: `http://www.ces.ncsu.edu/nreos/forest/pdf/WON/won26.pdf` [accessed 9 December 2006].

Möttönen, J. & Nummi, T. (2002), "Scots pine stem-curve predictions", 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. 186–190.

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.

Näsberg, M. (1985), "Mathematical programming models for optimal log bucking", Linköping Studies in Science and Technology, Dissertation No. 132, Department of Mathematics, Linköping University, Linköping, Sweden.

Newnham, R. M. (1992), "Variable-form taper functions for four Alberta tree species", *Can. J. For. Res.*, 22, 210–223.

Nummi, T. & Möttönen, J. (2004a), "Prediction of stem measurements of Scots pine", *J. Appl. Statist.*, 31, 105–114.

——— (2004b), "Estimation and prediction for low degree polynomial models under measurement errors with an application to forest harvesters", *Applied Statistics*, 53, 495–505.

Nummi, T., Sinha, B. K. & Koskela, L. (2005), "Statistical properties of the apportionment degree and alternative measures in bucking outcome", *Revista Investigacion Operaciónal*, 26, 259–267.

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.

Ponsse Oyj (2006), "Annual Report 2005", Ponsse Oyj, Kajaani, Finland.

Potthoff, R. F. & Roy, S. N. (1964), "A generalized multivariate analysis of variance model useful especially for growth curve problems", *Biometrika*, 51, 313–326.

Pulkki, R. (1997), "Cut-to-length, tree-length or full tree harvesting?", Central Woodlands, 1, n:o 3, 22–27, 37.

Stokes, B. J., Ashmore, C., Rawlings, C. L. & Sirois, D. L. (1989), "Glossary of terms used in timber harvesting and forest engineering", Gen. Tech. Rep. SO-73. New Orleans, LA: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station, 33 p.

Uusitalo, J. & Kivinen, V.-P. (2001), "Factors affecting the output distribution of logs with modern tree bucking optimizing system", in *Proceedings of the 2nd World Symposium on Logistics in Forest Sector*, Växjö, Sweden, 12–15 August, 2001, eds. K. Sjöström & L.-O. Rask, Växjö University School of Industrial Engineering and Timber Logistics, pp. 133–142.

Verbyla, A., Cullis, B. R., Kenward, M. G. & Welham, S. J. (1999), "The analysis of designed experiments and longitudinal data by using smoothing splines", *Appl. Statist.*, 48, 269–311.

Vuorenpää, T., Aaltonen, A., Imponen, V. & Lukkarinen, E. (1997), "Tukkijakauman ohjaus [Control of output log distributions]", Metsätehon raportti 38, Metsäteho Oy, Helsinki, Finland, 24 p. [In Finnish.]

Weijo, A. (2000), "Runkopankki Metsähallituksen puunhankinnan ohjauksen apuvälineenä [The stem bank as a control tool in the wood procurement of the Metsähallitus corporation]", Master's thesis, Faculty of Forestry, University of Joensuu, Joensuu, Finland. [In Finnish.]

Appendix A

Supplementary to Paper III

A.1 Approximating the Variance of the Price-weighted χ^2-statistic

The statements on the approximation of the variance of the $\chi^2(p^*)$ given on page 263 in the paper are justified by noting that

$$\begin{aligned} V(\chi^2(p^*)) &= V(\sum\sum p_{ij}^* z_{ij}^2) \approx V(\sum\sum{}' p_{ij}^* z_{ij}^2) \\ &\quad \text{(the term with the smallest } p_{ij}^* \ (<< \tfrac{1}{mn}) \text{ is taken out} \\ &\quad \text{leaving } nm-1 \text{ independent } \chi_1^2 \text{ terms)} \\ &= \sum\sum{}' {p_{ij}^*}^2 V(z_{ij}^2) = \sum\sum{}' {p_{ij}^*}^2 V(\chi_1^2) \\ &= \sum\sum{}' {p_{ij}^*}^2 \cdot 2 = 2\sum\sum{}' {p_{ij}^*}^2 \\ &\quad \text{(the term with the smallest } p_{ij}^* \text{ is again brought back)} \\ &\approx 2\sum\sum {p_{ij}^*}^2. \end{aligned}$$

Since

$$2nm{p_{min}^*}^2 = 2\sum\sum {p_{min}^*}^2 \leq 2\sum\sum {p_{ij}^*}^2 \leq 2\sum\sum {p_{max}^*}^2 = 2nm{p_{max}^*}^2,$$

one approximation is

$$V(\chi^2(p^*)) \approx \frac{2nm{p_{min}^*}^2 + 2nm{p_{max}^*}^2}{2} = nm({p_{min}^*}^2 + {p_{max}^*}^2).$$

However, a better approximation for the variance is achieved by noting that

$$2p_{min}^* = 2\sum\sum p_{ij}^* p_{min}^* \leq 2\sum\sum {p_{ij}^*}^2 \leq 2\sum\sum p_{ij}^* p_{max}^* = 2p_{max}^*,$$

and hence

$$V(\chi^2(p^*)) \approx \frac{2p_{min}^* + 2p_{max}^*}{2} = p_{min}^* + p_{max}^*.$$

A.2 Description of the Simulation Procedure

In the simulation study conducted in Section 3, new output matrices were created by randomly deleting a given portion of logs from the given target matrix. Letting $p\%$ be the portion of logs to be deleted and N the total number of logs in the target table, the simulation proceeded by repeating Steps 1-3 below $pN/100$ times.

1. Randomly select a log category out of the 25 in the target matrix.
2. The number of logs in the selected log category is reduced by one.
3. If the number of logs in the selected log category is zero, i.e. there are no logs in the log category, the log number will not be changed.

Note that due to Step 3, the eventual total number of deleted logs may not necessarily equal $pN/100$ for larger p.

A.3 Justification for the Use of χ^2-distribution

Assume $\mathbf{f} = (f_1, \ldots, f_{k+1})' \sim \text{Mult}(N; \boldsymbol{\pi})$, where $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_{k+1})' = \frac{1}{N}(\delta_1, \ldots, \delta_{k+1})'$. Then for the mean vector of $\mathbf{f}$ we get $\text{E}(\mathbf{f}) = N\boldsymbol{\pi}$ and the dispersion matrix we write as

$$\text{D}(\mathbf{f}) = N[\text{diag}(\boldsymbol{\pi}) - \boldsymbol{\pi}\boldsymbol{\pi}'] = N \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

where

$$\Sigma_{11} = \begin{pmatrix} \pi_1(1-\pi_1) & -\pi_1\pi_2 & \cdots & \pi_1\pi_k \\ -\pi_1\pi_2 & \pi_2(1-\pi_2) & \cdots & \pi_2\pi_k \\ \vdots & & \ddots & \vdots \\ -\pi_1\pi_k & -\pi_2\pi_k & \cdots & \pi_k(1-\pi_k) \end{pmatrix} = \text{diag}(\boldsymbol{\pi}^*) - \boldsymbol{\pi}^*\boldsymbol{\pi}^{*\prime}$$

and $\boldsymbol{\pi}^* = (\pi_1, \ldots, \pi_k)'$.

Since $\mathbf{1}'\mathbf{f} = N \Rightarrow \text{Var}(\mathbf{1}'\mathbf{f}) = 0 \Leftrightarrow \mathbf{1}'\text{D}(\mathbf{f})\mathbf{1} = 0$, $\text{D}(\mathbf{f})$ is singular. Hence we now restrict attention to only the first k f_i's, i.e. $\mathbf{f}^* = (f_1, \ldots, f_k)'$, for which $\text{E}(\mathbf{f}^*) = N\boldsymbol{\pi}^*$ and $\text{D}(\mathbf{f}^*) = N\Sigma_{11}$. Note that $\text{D}(\mathbf{f}^*)$ is p.d. It follows from the multivariate central limit theorem (MCLT, see e.g. Bilodeau & Brenner, 1999, pp. 78) that $\mathbf{f}^* \to \text{N}_k\left(N\boldsymbol{\pi}^*, \text{D}(\mathbf{f}^*)\right)$ and hence

$$(\mathbf{f}^* - N\boldsymbol{\pi}^*)' \left[\text{D}(\mathbf{f}^*)\right]^{-1} (\mathbf{f}^* - N\boldsymbol{\pi}^*) \sim \chi^2_k.$$

We may now write

$$\text{D}(\mathbf{f}^*) = N(\mathbf{A} - \boldsymbol{\alpha}\boldsymbol{\alpha}'),$$

where $\mathbf{A} = \text{diag}(\boldsymbol{\pi}^*)$ and $\boldsymbol{\alpha} = \boldsymbol{\pi}^*$. Then

$$\begin{aligned}
[\mathrm{D}(\mathbf{f}^*)]^{-1} &= \frac{1}{N}(\mathbf{A} - \boldsymbol{\alpha}\boldsymbol{\alpha}')^{-1} = \frac{1}{N}\left[\mathbf{A}^{-1} + \frac{\mathbf{A}^{-1}\boldsymbol{\alpha}\boldsymbol{\alpha}'\mathbf{A}^{-1}}{1 - \boldsymbol{\alpha}'\mathbf{A}\boldsymbol{\alpha}}\right] \\
&= \frac{1}{N}\left[\begin{pmatrix} \pi_1^{-1} & & \\ & \ddots & \\ & & \pi_k^{-1} \end{pmatrix} + \frac{\mathbf{1}\mathbf{1}'}{1 - \sum_{i=1}^k \pi_i}\right] \\
&= \left[\begin{pmatrix} \delta_1^{-1} & & \\ & \ddots & \\ & & \delta_k^{-1} \end{pmatrix} + \frac{\mathbf{1}\mathbf{1}'}{N - \sum_{i=1}^k \delta_i}\right]
\end{aligned}$$

and

$$\begin{aligned}
&(\mathbf{f}^* - N\boldsymbol{\pi}^*)' \left[\mathrm{D}(\mathbf{f}^*)\right]^{-1} (\mathbf{f}^* - N\boldsymbol{\pi}^*) \\
&= \left[(f_1 - \delta_1, \ldots, f_k - \delta_k)\left\{\begin{pmatrix} \delta_1^{-1} & & \\ & \ddots & \\ & & \delta_k^{-1} \end{pmatrix} + \frac{\mathbf{1}\mathbf{1}'}{N - \sum_{i=1}^k \delta_i}\right\}\begin{pmatrix} f_1 - \delta_1 \\ \vdots \\ f_k - \delta_k \end{pmatrix}\right] \\
&= \sum_{i=1}^k \frac{(f_i - \delta_i)^2}{\delta_i} + \frac{\sum_{i=1}^k (f_i - \delta_i) \sum_{i=1}^k (f_i - \delta_i)}{N - \sum_{i=1}^k \delta_i} \\
&= \sum_{i=1}^k \frac{(f_i - \delta_i)^2}{\delta_i} + \frac{\left(\sum_{i=1}^k (f_i - \delta_i)\right)^2}{N - \sum_{i=1}^k \delta_i} \\
&= \sum_{i=1}^k \frac{(f_i - \delta_i)^2}{\delta_i} + \frac{(f_{k+1} - \delta_{k+1})^2}{\delta_{k+1}} = \sum_{i=1}^{k+1} \frac{(f_i - \delta_i)^2}{\delta_i} = \chi^2,
\end{aligned}$$

and hence $\chi^2 \sim \chi^2_k$.

A.4 Justification for the Use of the Weighted χ^2-distribution

Now we justify the use of the weighted χ^2-distribution for a measure with positive weights. This measure is denoted by $\chi^2(w)$. We use the same notations as in Appendix A.3 and assume $\mathbf{f} = (f_1, \ldots, f_{k+1})' \sim \text{Mult}(N; \boldsymbol{\pi})$, where $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_{k+1})' = \frac{1}{N}(\delta_1, \ldots, \delta_{k+1})'$. Then,

$$\chi^2(w) = \sum_{i=1}^{k+1} \frac{w_i (f_i - \delta_i)^2}{\delta_i}.$$

The $(k+1)$th term in the $\chi^2(w)$ can be written as

$$\begin{aligned}
w_{k+1}\frac{(f_{k+1} - \delta_{k+1})^2}{\delta_{k+1}} &= w_{k+1}\frac{\left(\sum_{i=1}^k (f_i - \delta_i)\right)^2}{\delta_{k+1}} = \frac{w_{k+1}}{\delta_{k+1}}(\mathbf{f}^* - \boldsymbol{\delta}^*)'\mathbf{1}\mathbf{1}'(\mathbf{f}^* - \boldsymbol{\delta}^*) \\
&= (\mathbf{f}^* - \boldsymbol{\delta}^*)'\mathbf{A}^*(\mathbf{f}^* - \boldsymbol{\delta}^*)
\end{aligned}$$

where $\mathbf{f}^* = (f_1, \ldots, f_k)'$, $\boldsymbol{\delta}^* = (\delta_1, \ldots, \delta_k)'$ and $\mathbf{A}^* = \frac{w_{k+1}}{\delta_{k+1}}\mathbf{1}\mathbf{1}'$. The other terms in the $\chi^2(w)$ can be written as

$$(\mathbf{f}^* - \boldsymbol{\delta}^*)'\mathbf{B}(\mathbf{f}^* - \boldsymbol{\delta}^*),$$

where $\mathbf{B} = \text{diag}(\frac{w_1}{\delta_1}, \ldots, \frac{w_k}{\delta_k})$. Hence

$$\chi^2(w) = (\mathbf{f}^* - \boldsymbol{\delta}^*)'(\mathbf{A}^* + \mathbf{B})(\mathbf{f}^* - \boldsymbol{\delta}^*) = \mathbf{x}'(\mathbf{A}^* + \mathbf{B})\mathbf{x},$$

where $\mathbf{x} = (\mathbf{f}^* - \boldsymbol{\delta}^*)$.

Since $\mathbf{f}^* \rightarrow \text{N}_k(\boldsymbol{\delta}^*, \text{D}(\mathbf{f}^*))$, then $\mathbf{x} = (\mathbf{f}^* - \boldsymbol{\delta}^*) \rightarrow \text{N}_k(\mathbf{0}, \text{D}(\mathbf{x}))$, where $\text{D}(\mathbf{x}) = \text{D}(\mathbf{f}^*) = N(\mathbf{A} - \boldsymbol{\alpha}\boldsymbol{\alpha}')$ is a positive definite matrix and $\mathbf{A}$ and $\boldsymbol{\alpha}$ are as defined in Appendix A.3 above. Let $\mathbf{z} = \text{D}(\mathbf{x})^{-\frac{1}{2}}\mathbf{x}$. Then, $\mathbf{z} \rightarrow \text{N}_k(\mathbf{0}, \mathbf{I})$. We now look at the distribution of

$$\mathbf{x}'(\mathbf{A}^* + \mathbf{B})\mathbf{x} = \left(\left[\text{D}(\mathbf{x})^{\frac{1}{2}}\right]\mathbf{z}\right)'(\mathbf{A}^* + \mathbf{B})\left(\left[\text{D}(\mathbf{x})^{\frac{1}{2}}\right]\mathbf{z}\right) = \mathbf{z}'\mathbf{C}\mathbf{z},$$

where

$$\begin{aligned}
\mathbf{C} &= \left[\text{D}(\mathbf{x})^{\frac{1}{2}}\right](\mathbf{A}^* + \mathbf{B})\left[\text{D}(\mathbf{x})^{\frac{1}{2}}\right] \\
&= \left[\text{D}(\mathbf{x})^{\frac{1}{2}}\right]\mathbf{A}^*\left[D(\mathbf{x})^{\frac{1}{2}}\right] + \left[\text{D}(\mathbf{x})^{\frac{1}{2}}\right]\mathbf{B}\left[\text{D}(\mathbf{x})^{\frac{1}{2}}\right] \\
&= \left[\text{D}(\mathbf{x})^{\frac{1}{2}}\right][\tfrac{x_{k+1}}{\delta_{k+1}}\mathbf{1}\mathbf{1}']\left[\text{D}(\mathbf{x})^{\frac{1}{2}}\right] + \left[\text{D}(\mathbf{x})^{\frac{1}{2}}\right]\text{diag}\left(\frac{w_1}{\delta_1}, \ldots, \frac{w_k}{\delta_k}\right)\left[\text{D}(\mathbf{x})^{\frac{1}{2}}\right].
\end{aligned}$$

Since $\mathbf{A}^*$ is symmetric and $\mathbf{B}$ diagonal, then clearly $\mathbf{A}^* + \mathbf{B}$ and also $\mathbf{C}$ are symmetric. According to eigenvalue decomposition there now exists a diagonal matrix $\boldsymbol{\Lambda}$ and an orthogonal matrix $\mathbf{T}$ such that $\mathbf{C} = \mathbf{T}'\boldsymbol{\Lambda}\mathbf{T}$. Letting $\lambda_1, \ldots, \lambda_k$ represent the diagonal elements of $\boldsymbol{\Lambda}$ and $y_1, \ldots, y_k$ the elements of vector $\mathbf{y} = \mathbf{T}\mathbf{z}$, we may find that

$$\chi^2(w) = \mathbf{z}'\mathbf{C}\mathbf{z} = \mathbf{z}'\mathbf{T}'\boldsymbol{\Lambda}\mathbf{T}\mathbf{z} = (\mathbf{T}\mathbf{z})'\boldsymbol{\Lambda}(\mathbf{T}\mathbf{z}) = \mathbf{y}'\boldsymbol{\Lambda}\mathbf{y} = \sum_{i=1}^{k}\lambda_i y_i^2,$$

where $\mathbf{y} = \mathbf{T}\mathbf{z} \rightarrow \text{N}_k(\mathbf{0}, \mathbf{T}\mathbf{I}\mathbf{T}') = \text{N}_k(\mathbf{0}, \mathbf{I})$, i.e. $y_i \sim$ *i.i.d.* $\text{N}(0, 1)$, $i = 1, \ldots, k$. This shows that $\chi^2(w)$ is a weighted sum of k independent χ_1^2 distributed terms. Since the weights are positive, we may approximate the distribution of $\chi^2(w)$ by $a\chi_b^2$ for some suitable choices of a and b.

Appendix B

Supplementary to Paper VI

In paper V we adopted Dirichlet distribution to describe the joint distribution of the random components in the output matrix and aimed at choosing parameters of the distribution such that the Apportionment Index is maximized in the averaged sense. In the case of two log categories, we proposed specifying an upper bound for the variance of the random outputs to find a unique solution. However, the extension of the method to the case of three log categories turned out to be analytically intractable. A technique was then proposed to tackle the problem of specifying the parameter values when the target matrix is given and an upper bound to the largest of the variances of the random outputs is specified. However, the appropriateness of the technique in terms of maximizing the $E(AI)$ was not studied in the paper. We now take up some computations on the suitability of the technique.

We examine the suitability for the three settings in Example 2.1 by computing the probabilities $\Gamma_1 = \int_0^{\theta_1} B(\alpha_1, \alpha_2 + \alpha_3)$, $\Gamma_2 = \int_0^{\theta_2} B(\alpha_2, \alpha_1 + \alpha_3)$ and $\Gamma_3 = \int_0^{\theta_3} B(\alpha_3, \alpha_1 + \alpha_2)$. In an ideal case θ_i, $i = 1, 2, 3$, should serve as a median of the marginal distribution of the respective X_i. Hence, the closer the values of Γ_i's are to 0.5, the closer the solution is to the optimal.

Table B.1 shows that for smaller V_0 the parameter values provided by the proposed technique are relatively close to optimal in all three settings. However, as the variance is increased the solutions start to deviate from the optimal in all three settings. A closer look at the results shows that all Γ's exceed 0.5. This indicates that all three marginal distributions which are specified by the parameters provided by the proposed technique possess slighly heavy left tails.

Earlier we proposed choosing the parameter values in the ratio of the known θ's, i.e. by setting $\alpha_1 = \alpha$, $\alpha_2 = \frac{\theta_2}{\theta_1}\alpha = a\alpha$ and $\alpha_3 = \frac{\theta_3}{\theta_1}\alpha = b\alpha$ and solving for α for a specified V_0, which corresponds to the highest marginal variance of the X_is, namely that of X_3. The above calculations indicate, however, that we might possibly achieve a better solution by "fine-tuning" the proposed technique. Let us set $\alpha_1 = \alpha$, $\alpha_2 = \left(\frac{\theta_2}{\theta_1}\right)^{\gamma}\alpha = a^{\gamma}\alpha$ and $\alpha_3 = \left(\frac{\theta_3}{\theta_1}\right)^{\gamma}\alpha = b^{\gamma}\alpha$, where γ is some fixed quantity in the interval $(0, 1)$ and α is the parameter specified by the given variance constraint. By an appropriate choice of γ, the forms of the marginal distributions can be fine-tuned so as to shift the medians closer to the respective target values, i.e. to improve the choice in the sense of optimality. For a given target and a specified V_0, the

choice of γ could be made, for example, by solving

$$\frac{1}{3}\left(\Gamma_1 + \Gamma_2 + \Gamma_3\right) \approx 0.5.$$

As discussed in the paper, the problem of specifying the model parameters by maximizing the expected value of the AI is a complex task and leads to highly complicated analytical computations even in the case of only three locations. However, as seen in Table B.1 below, more or less satisfactory results can be achieved for small V_0. Possibly, satisfactory results could also be achieved for larger V_0 by developing the idea introduced above. Thereafter, the extension of the technique to a more general case, i.e. to some moderate number of locations (> 3), could be justified. However, further research is required to determine whether significant improvements are achieved by the modifications presented above, or otherwise. Our initial computations towards incorporating the idea of introducing the scaling factor γ, mentioned above, do not indicate any significant improvements. We leave study in this direction as a topic for future research.

Table B.1: Computations assessing the optimality of the solutions in Example 2.1 in Paper VI.

	$\theta_1 = 0.17$ $\theta_2 = 0.38$ $\theta_3 = 0.45$				$\theta_1 = 0.10$ $\theta_2 = 0.45$ $\theta_3 = 0.45$			$\theta_1 = 1/3$ $\theta_2 = 1/3$ $\theta_3 = 1/3$	
V_0	α	Γ_1	Γ_2	Γ_3	α	Γ_1	$\Gamma_2 = \Gamma_3$	α	$\Gamma_1 = \Gamma_2 = \Gamma_3$
0.01	4.04	0.548	0.514	0.506	2.38	0.573	0.506	7.07	0.521
0.02	1.93	0.570	0.520	0.508	1.14	0.606	0.508	3.37	0.530
0.03	1.23	0.588	0.525	0.510	0.73	0.631	0.510	2.14	0.538
0.04	0.88	0.604	0.530	0.512	0.52	0.654	0.512	1.52	0.545
0.05	0.67	0.618	0.534	0.514	0.47	0.661	0.513	1.15	0.552
0.075	0.39	0.652	0.544	0.518	0.23	0.717	0.518	0.65	0.569
0.10	0.25	0.682	0.555	0.522	0.15	0.751	0.522	0.41	0.584
0.125	0.17	0.708	0.565	0.527	0.098	0.784	0.527	0.26	0.601
0.15	0.11	0.737	0.576	0.531	0.065	0.811	0.531	0.16	0.618
0.175	0.070	0.762	0.587	0.536	0.041	0.837	0.536	0.090	0.634
0.20	0.040	0.786	0.598	0.541	0.024	0.860	0.541	0.037	0.651

Appendix C

Errata to the Original Papers

Paper I

• Page 2914

– Equation 8 Printed: $\hat{s} = \mu_1 + \frac{\sigma_{01}}{\sigma_0^2}(\mathrm{DBH} - \mu_0)$

Should read: $\hat{s} = \mu_0 + \frac{\sigma_{01}}{\sigma_1^2}(\mathrm{DBH} - \mu_1)$

Paper III

• Page 260

– Line 7 Typing error. Should read: "... may yield very undesirable..."

– Line 11 Typing error. Should read: "Similarly we can..."

– Line 18 Printed: $^{*}o_{ij}^{*} = \frac{o_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{m} o_{ij}}$

Should read: $\mathrm{o}_{\mathrm{ij}}^{*} = \frac{\mathrm{o}_{\mathrm{ij}}}{\sum_{\mathrm{i=1}}^{\mathrm{n}}\sum_{\mathrm{j=1}}^{\mathrm{m}} \mathrm{o}_{\mathrm{ij}}}$

– Line 27 Printed: $\mathrm{A_p} = \sum_{i=1}^{n}\sum_{j=1}^{m}\delta_{ij}{}^{*}p_{ij}^{*}$,

Should read: $\mathrm{A_p} = \sum_{\mathrm{i=1}}^{\mathrm{n}}\sum_{\mathrm{j=1}}^{\mathrm{m}}\delta_{\mathrm{ij}}\mathrm{p}_{\mathrm{ij}}^{*}$,

– Line 28 Printed: $^{*}p_{ij}^{*} = \frac{1}{\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}}$

Should read: $\mathrm{p}_{\mathrm{ij}}^{*} = \frac{\mathrm{p}_{\mathrm{ij}}}{\sum_{\mathrm{i=1}}^{\mathrm{n}}\sum_{\mathrm{j=1}}^{\mathrm{m}} \mathrm{p}_{\mathrm{ij}}}$

• Page 261

– Line 3 All the characters in the two equations should be written in non-cursive style.

- Page 262

– Line 1 — The epsilon in the equation should follow the style used elsewhere in the paper.

– Line 3 — t_{ij} should be written in non-cursive style in both appearances on the line.

– Lines 5–6 — All the characters in the equation should be written in non-cursive style.

– Lines 10 — All the characters in the equation should be written in non-cursive style.

– Lines 3–12 — There appear slightly different summation notations on the page. The notations $\sum^{'}$ and $\sum^{''}$ correspond to $\sum^{/}$ and $\sum^{//}$, respectively. Unique summation notations should be used for both cases.

- Page 263

– Line 11 — Printed: $-\ \mathrm{V}(\chi^2(\mathrm{p}^*)) \approx 2\mathrm{nm}(\mathrm{p}^{*2}_{\mathrm{min}} + \mathrm{p}^{*2}_{\mathrm{max}})/2 = nm(p^{*2}_{\mathrm{min}} + p^{*2}_{\mathrm{max}})''$

Should read: $\mathrm{V}(\chi^2(\mathrm{p}^*)) \approx 2\mathrm{nm}(\mathrm{p}^{*2}_{\mathrm{min}} + \mathrm{p}^{*2}_{\mathrm{max}})/2 = \mathrm{nm}(\mathrm{p}^{*2}_{\mathrm{min}} + \mathrm{p}^{*2}_{\mathrm{max}})$

– Line 19 — Printed: $\hat{\mathrm{b}} = 2/\mathrm{t}$

Should read: $\hat{\mathrm{b}} = 2/\mathrm{T}$

- Page 264

– Line 23 — Typing error. Should read: "... is approximately the same..."

– Line 27 — Typing error. Should read: "... is approximately 0.89..."

- Page 266

– Figure 2 — The title of the bottom figure is missing. The figure should be titled: "(c) Results of $\mathrm{A_s}(\mathrm{p}^*_2)$." The Apportionment Index notations in the other figure titles should be written in non-cursive style.

- Page 267

– Figure 3 — The title of the bottom figure is missing. The figure should be titled: "(c) Associated p-values of $\chi^2(\mathrm{p}^*_2)$." The notation p^*_1 in the figure title (b) should be written in non-cursive style.

Paper VI

- Page 169

– Lines 45–46 — Missing capital letters. Should read: "In Section 2.1 we take the analysis of the two log classes as a starting-point. In Section 2.2 we extend the analysis to the case of three locations and finally in Section 3..."

– Footnote — Printed: laura.koskela@uta.fi*

Should read: laura.koskela@uta.fi

– Footnote — Printed: tan@uta.fi*

Should read: tan@uta.fi

• Page 170

– Line 5 — Printed: $\mathrm{AI} = \sum_{\mathrm{i}} \sum_{\mathrm{j}} \min\{\mathrm{X_{ij}} \ . \ \theta_{\mathrm{ij}}\}$.

(Note that θ could not be written above in non-cursive style for technical reasons.)

Should read: $\mathrm{AI} = \sum_{\mathrm{i}} \sum_{\mathrm{j}} \min\{\mathrm{X_{ij}}, \theta_{\mathrm{ij}}\}$,

(Note that θ could not be written above in non-cursive style for technical reasons.)

• Page 171

– Line 3 — Extra space. Should read: "quality, say α, in view..."

• Page 172

– Line 12 — Printed: $\mathrm{D}(\alpha_1, \alpha_2, \alpha_3) = \Gamma(\alpha_1, \alpha_2, \alpha_3)/\ldots$

Should read: $\mathrm{D}(\alpha_1, \alpha_2, \alpha_3) = \Gamma(\alpha_1 + \alpha_2 + \alpha_3)/\ldots$

• Page 173

– Line 3 — Missing dot after Equation 13.

• Page 176

– Table 5 — Incorrect equation in the title of the table.

Printed: $\mathrm{RR(i,j)} = \frac{|\mathrm{Q_i} - \mathrm{Q_j}|}{\frac{1}{2}(\mathrm{Q_i} - \mathrm{Q_j})} \times 100\ \%$

Should read: $\mathrm{RR(i,j)} = \frac{|\mathrm{Q_i} - \mathrm{Q_j}|}{\frac{1}{2}(\mathrm{Q_i} + \mathrm{Q_j})} \times 100\ \%$

– Line 16 — Typing error. Should read: "Fördelningsaptering - ett sätt att tillgodose sågverksönskemål.""

– Line 23 — Typing error. Should read: "apteeraustuloksen arvioinnissa..."

– Line 24 — Typing error. Should read: "1/2000: 59–61."

– Line 27 — Missing comma. Should read: "Skogforsk 537, 87–95), Växjö, Sweden."

• Page 177

– Lines 5–6 — Extra colon and extra comma. Should read: "...4^{th} ed. Springer. New York."

STATISTICAL PROPERTIES OF THE APPORTIONMENT DEGREE AND ALTERNATIVE MEASURES IN BUCKING OUTCOME

Tapio Nummi*, Bikas K. Sinha** and Laura Koskela*
*Department of Mathematics, Statistics and Philosophy, Statistics Unit, University of Tampere, Finland
**Indian Statistical Institute, Kolkata, India

ABSTRACT
In the harvesting technique prevailing in Scandinavia, tree stems are converted into smaller logs immediately at harvest. Modern sawmills attempt to operate according to customers' special needs rather than only minimize the production costs. Since the annual production of saw timber in Scandinavia is in tens of millions of cubic meters, proper measuring of the goodness of the bucking outcome is of crucial importance. The outcome of the bucking operation can be considered as a multidimensional table of tree species, quality grades, prices and length and diameters classes. The prevailing method to measure the outcome is the so-called apportionment degree, which is calculated from the relative portions of the observed and target tables. However, this measure has severe drawbacks. E.g. it gives the same weight for each log class. Therefore, for example, the effect of the shape of the distributions is completely ignored. In this study we present some basic results of the statistical properties of the apportionment degree and present some alternative means to measure the bucking outcome. Also a simulation study is carried out to illustrate the relative performance of the measures presented.

Key words: Apportionment degree, forest harvesting, frequency Chi-Square, simulation.

MSC: 62P12

RESUMEN
El propósito de este trabajo es comenzar el análisis estadístico del índice del prorrateo, esta es una medida usada en el contexto de la tala en bosques para evaluar el ajuste entre la demanda y distribución del suministro de la leña. Ha habido algunos esfuerzos por entender este índice, pero una base teórica seria todavía falta. Nosotros discutimos brevemente la literatura existente y procedemos a investigar las propiedades del índice desde un punto de vista distribucional. Este es fundamentalmente un artículo exploratorio y nosotros sólo enfocamos los casos de dos y tres clases de leña, es decir, locaciones. En el caso de dos clases usamos la distribución beta para las variables aleatorias relativas a la salida (output) del rendimiento; en tres locaciones al azar se asume que los rendimientos relativos siguen la distribución de Dirichlet singular. Usando esta formulación es posible entender las propiedades estadísticas del índice del prorrateo.

1. INTRODUCTION

The general objective in harvesting is to maximize the value of the timber obtained for further processing. Optimization of harvesting requires that several phases in a production chain are successfully combined. In the harvesting technique prevailing in Scandinavia, tree stems are converted into smaller logs immediately at harvest. High-class measuring and computing equipment have been developed, making possible computer-based optimization of crosscutting in harvesters. In modern harvesters tree stems are run in sequence through the measuring equipment and simultaneously the harvester's computer receives the length and diameter data from sensors. If the whole stem is measured before crosscutting we may apply the techniques discussed e.g. in Näsberg (1985) to find the optimal cutting patterns on the stem. However, in practice the first cutting decisions have to be made under incomplete stem information and we must compensate the unknown part of the stem by predictions (see e.g Liski and Nummi (1995)).

An admissible cutting pattern is a set of cutting points $0 = x_1 < x_2 < \ldots < x_R$ such that the length of the rth log

$$l_r = x_r - x_{r-1} \in [l_{min}, l_{max}] \text{ and } d(x_r) \geq d_T > 0$$

for r = 2, 3,..., R, where $x_1 = 0$ is at the butt of a tree, l_{min} is the minimum and l_{max} the maximum length of a log and d_T is the minimum acceptable log diameter. Marking for bucking is the problem of converting a single tree stem into logs in such a way that the total stem value (price, volume etc.) for logs is maximized (see Näsberg 1985, Chapter 3).

We can classify a log with a small end diameter d(x) and length l (index r dropped) to one of the m × n classes according to the following classification

$$d_i \le d(x) < d_{i+1} \quad \text{and} \quad l_j \le l < l_{j+1},$$

where d_i, i = 1,…,n and l_j, j = 1,…,m are given diameter and length limits. Then we may for example specify the price of each diameter and length combination d_i, l_j of logs. Denote these as the m × n price matrix **P**, where the element p_{ij} of **P** is the price of the log at log class d_i, l_j. However, it is well known that optimization of price only may yield very undesiderable log distributions from the sawmills point of view. Nowadays sawmills aim to operate more on customers special needs rather than maximizing price or minimizing the production costs only. In fact we may have simultaneously many targets. Especially we may have a matrix of frequencies jointly with a matrix of prices. We may define the target amount of logs for each diameter and length combinations. Denote these as the m × n target matrix **T**. Similarily we can classify the outcome of the actual bucking operation to elements of the m × n frequency matrix **O**. Then the measures studied here are simple functions of the actual output **O**, the target **T** and the log prices **P**.

2. MEASURING THE BUCKING OUTCOME

2.1. The Apportionment degree

The so-called apportionment degree widely used in harvesting is defined for a fixed quality class as follows

$$A = \left(1 - 0.5 \times \sum_{i=1}^{n} \sum_{j=1}^{m} \left| o_{ij}^* - t_{ij}^* \right| \right),$$

where $^{*}o_{ij}^{*} = \dfrac{o_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{m} o_{ij}}$ and $t_{ij}^* = \dfrac{t_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{m} t_{ij}}$, where o_{ij} and t_{ij} are elements of the outcome and target matrices, respectively. With this measure we can compare the relative proportions of the output and target tables. The apportionment degree A gives a value between 0 to 1, where the value A = 1 corresponds to perfect match of the tables. After some simple manipulations we can show that A can be rewritten as

$$A = \sum_{i=1}^{n} \sum_{j=1}^{m} \delta_{ij}, \tag{1}$$

where $\delta_{ij} = \min(o_{ij}^*, t_{ij}^*)$. This measure was first introduced by the swedish mathematician Bergstrand in the mid 1980s, when first steps in developing automatic bucking systems were taken. However very little is known of the statistical properties of the apportionment degree A.

It is easy to give a price-weighted version of A. Then we compute

$$A_p = \sum_{i=1}^{n} \sum_{j=1}^{m} \delta_{ij}{}^{*} p_{ij}^{*},$$

where $^{*}p_{ij}^{*} = \dfrac{1}{\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}}$ is the relative price of the log at given log class. However, A_p is no longer in the same magnitude as A. We first note that

$$\rho_{\delta,p^*} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} \delta_{ij} p_{ij}^* - \dfrac{A}{nm}}{\left[\left(\sum_{i=1}^{n}\sum_{j=1}^{m} \delta_{ij}^2 - \dfrac{A^2}{nm} \right) \left(\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{*2} - \dfrac{1}{nm} \right) \right]^{0.5}} \le 1$$

and, hence,

$$A_p = \sum_{i=1}^{n}\sum_{j=1}^{m}\delta_{ij}p_{ij}^{*} \le \frac{A}{nm} + \sqrt{\left(\sum_{i=1}^{n}\sum_{j=1}^{m}\delta_{ij}^{2} - \frac{A^2}{nm}\right)\left(\sum_{i=1}^{n}\sum_{j=1}^{m}\delta_{ij}^{*2} - \frac{1}{nm}\right)}.$$

We denote $\sigma_\delta = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m}\delta_{ij}^{2} - \frac{A^2}{nm}}$ and $\sigma_{p^*} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m}p_{ij}^{*2} - \frac{1}{nm}}$. Then it is easy to see that the values of A_p are at the interval

$$A_p \in \left[0, \frac{A}{nm} + \sigma_\delta \sigma_{p^*}\right].$$

Then instead of A_p we can calculate

$$A_s = \frac{nmA_p}{A + nm\sigma_\delta\sigma_{p^*}},$$

which is of the same magnitude as the original A. Note that our choice $\rho_{\delta,p^*} = 1$ generally overestimates the true value of ρ_{δ,p^*}. This may imply that on the average the scaled statistic A_s underestimates the true apportionment degree.

It is now easy to make some observations concerning A_s. First if δ and p* are independent we note that $A_s = nmA_p/A$. Similarly if $\sigma_\delta \approx 0$ or $\sigma_{p^*} \approx 0$ we note that in both cases $A_s = nmA_p/A$. These correspond to situations where price is approximately uniform or the disparity between demand and supply is more or less uniform, respectively.

2.2. Analysis with standard statistical measures

One of the most common measures to test the fit between two distributions is the χ^2-test. By using our notations this statistic is defined as

$$\chi^2 = \sum_{i=1}^{n}\sum_{j=1}^{m}(o_{ij} - t_{ij})^2 / t_{ij},$$

which under certain conditions follows the χ^2-distribution with nm - 1 degrees of freedom. Note that here we use nm instead of nm - 1 degrees of freedom as an approximation since nm is in practical situations appropriately large. A price-weighted version of the statistic can be written as

$$\chi^2(p^*) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}p_{ij}(o_{ij} - t_{ij})^2 / t_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{m}p_{ij}} \qquad (2)$$

It can be shown see e.g. Rao (1973) that the distributions of this statistic can be approximated by weighted χ^2-distribution.

2.2.1. Relation to the apportionment degree in the unweighted case

We first write A as in (1), and

$$\chi^2 = \sum_{i=1}^{n}\sum_{j=1}^{m}\left(o_{ij}^2 / t_{ij} - \frac{N}{nm}\right), \qquad (3)$$

where we assume $N = \sum_{i=1}^{n}\sum_{j=1}^{m}t_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{m}o_{ij}$. Then define

$$o_{ij} = t_{ij} + \varepsilon_{ij}, \ \forall i,j$$

where we assume $\sum_{i=1}^{n}\sum_{j=1}^{m} \in_{ij} = 0$,

$$\sum^{'} = \sum_{\{(i,j)\,|\,o_{ij} < t_{ij}\}} \text{ and } \sum^{''} = \sum_{\{(i,j)\,|\,o_{ij} \geq t_{ij}\}} .$$

By using these notations we write the χ^2-statistic in (3) as

$$\chi^2 = \sum^{I}\left[\frac{(t_{ij} + \in_{ij})^2}{t_{ij}} - \frac{N}{nm} \right] + \sum^{''} \left[\frac{(t_{ij} + \in_{ij})^2}{t_{ij}} - \frac{N}{nm} \right]$$

$$= \sum^{I}\left[(t_{ij} + 2 \in_{ij}) + (\in_{ij}^2 / t_{ij}) - \frac{N}{nm} \right] + \sum^{''} \left[(t_{ij} + 2 \in_{ij}) + (\in_{ij}^2 / t_{ij}) - \frac{N}{nm} \right].$$

Now we can write

$$\chi^2 = \sum_{i=1}^{n}\sum_{j=1}^{m} t_{ij} + \sum^{I}\left(\in_{ij}^2 / t_{ij} - \frac{N}{nm} \right) + \sum^{II}\left(\in_{ij}^2 / t_{ij} - \frac{N}{nm} \right).$$

This can be further written as

$$\chi^2 = N(A - 1) - \sum^{I} \in_{ij} + \sum_{i=1}^{n}\sum_{j=1}^{m} \in_{ij}^2 / t_{ij} ,$$

where we note that in our notations

$$A = \frac{1}{N}\left(\sum_{i=1}^{n}\sum_{j=1}^{m} t_{ij} + \sum^{I} \in_{ij} \right).$$

2.2.2. Distribution of the weighted χ^2-statistic

The price-weighted χ^2-statistic in (2) can be written as a sum

$$\chi^2(p^*) = \sum_{i=1}^{n}\sum_{j=1}^{m} u_{ij}^2 ,$$

where

$$u_{ij} = \sqrt{p_{ij}^*}\, \frac{o_{ij} - t_{ij}}{\sqrt{t_{ij}}} .$$

It is easy to see that

$$u_{ij} \sim N(0,\ p_{ij}^*)$$

and

$$\frac{1}{\sqrt{p_{ij}^*}} u_{ij} = z_{ij} \sim N(0, 1).$$

Now $\chi^2(p^*) = \sum_{i=1}^{n}\sum_{j=1}^{m} u_{ij}^2 = \sum_{i=1}^{n}\sum_{j=1}^{m} z_{ij}^2 p_{ij}^*$ and hence the distribution of $\chi^2(p^*)$ is a weighted sum of linearly restricted χ_1^2 - variables . The distribution of $\chi^2(p^*)$ can be approximated by

$$\chi^2(p^*) \approx a\chi_b^2. \quad (4)$$

Now a and b can be solved from the first two moments of $\chi^2(p^*)$. The expected value is

$$E(\chi^2(p^*)) = 1.$$

For the variance we first note that the weights lie between the values

$$p^*_{min} \le p^*_{ij} \le p^*_{max}, \ \forall i,j.$$

Then we note that

$$p^*_{min}\chi^2_{nm} < \chi^2(p^*) < p^*_{max}\chi^2_{nm},$$

and therefore the variance is bounded by

$$2nmp^{*2}_{min} < V(\chi^2(p^*)) < 2nmp^{*2}_{max}.$$

Then one approximation would be

- $V(\chi^2(p^*)) \approx 2nm(p^{*2}_{min} + p^{*2}_{max})/2 = nm(p^{*2}_{min} + p^{*2}_{max})$."

We can now solve for a and b by equating the mean and variance of both sides in (4):

$$ab = 1$$

and

$$2a^2b = nm(p^{*2}_{min} + p^{*2}_{max}) = T, \text{ say.}$$

It follows that

$$\hat{a} = T/2$$

and

$$\hat{b} = 2/t,$$

where $\hat{a}$ and $\hat{b}$ are estimates of multiplier and degrees of freedom of the approximation (4), respectively.

2.2.3. A computational example

Assume that the target matrix is given in the Table 1. Then, for example, the target of the length of 430 cm and the top diameter of 160 mm logs is 28 objects. Assume that the actual output matrix is given in the Table 2. In fact the output matrix is obtained from the target matrix by randomly dropping 15 percent of logs from the target table.

Table 1. Target matrix.

Top diam (mm)	length (cm)					Total
	430	460	490	520	550	
160	28	16	58	45	45	192
200	37	17	65	45	37	201
240	17	49	37	44	55	202
280	22	39	39	44	59	203
340	19	30	47	54	52	202
Total	123	151	246	232	248	1000

Table 2. Output matrix.

Top diam (mm)	length (cm)					Total
	430	460	490	520	550	
160	23	12	52	39	41	167
200	33	11	61	39	27	171
240	12	39	30	38	49	168
280	14	34	33	36	56	173
340	8	30	42	47	44	171
Total	90	126	218	199	217	850

Table 3. Price matrix 1.

Top diam (mm)	length (cm)					Total
	430	460	490	520	550	
160	100	103	105	108	109	525
200	124	128	130	134	135	651
240	144	148	151	156	157	756
280	156	161	164	168	170	819
340	160	165	168	173	174	840
Total	684	705	718	739	745	3591

Next we investigate the fit between these two tables. The apportionment degree of these tables is A = 0.963. Since the value of the statistic is very close to 1 the fit between the two matrices is very good. The ordinary χ^2-test statistic gives the associated p-value 0.176. This comparison also shows that the fit between the target and the observed matrices is very good.

We may also specify the price of each diameter and length combination of logs. Here we use two price tables denoted by P_1 and P_2. The price matrix 1 is given in the Table 3 and P_2 is simply the matrix transpose of P_1. The scaled price-weighted apportionment degrees are $A_s(p_1^*) = 0.947$ and $A_s(p_2^*) = 0.982$, respectively. It is easy to see that also in this case the fit is very good, however a slightly better fit is obtained when the price matrix P_2 is used.

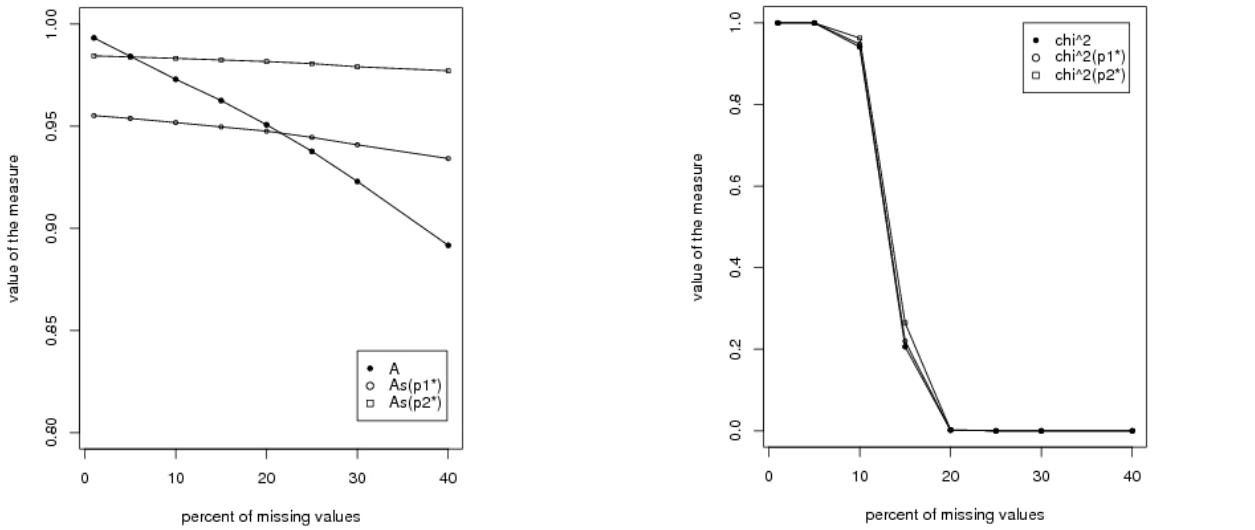
The price-weighted versions $\chi^2(p_1^*)$ and $\chi^2(p_2^*)$ gave the associated p-values 0.152 and 0.242, respectively. This also indicates a slightly better fit attained when using the price matrix P_2.

3. A SIMULATION STUDY

In this section we conduct a simulation study to investigate the performance of the apportionment degree, χ^2-test statistic and their price-weighted versions to measure the fit between target and output matrices. We take the target matrix in the Table 1 as a starting point. Next we randomly deleted 1%, 5%, 10%, 15%, 20%, 25% 30% and 40% in turn of the logs in the target table and this experiment is repeated 100 times at each percentage point. At each point the values of the apportionment degree, the χ^2-statistic and their price-weighted versions were calculated. For χ^2-statistics also the associated p-values were calculated.

Then mean curves of A, $A_s(p_1^*)$ and $A_s(p_2^*)$ are given in Figure 1a and mean curves of the p-values of χ^2, $\chi^2(p_1^*)$ and $\chi^2(p_2^*)$ are given in Figure 1b. From Figure 1a we observe that the average performance of the Apportionment degree A and its price-weighted versions $A_s(p_1^*)$ and $A_s(p_2^*)$ are approximately linear as a function of randomly generated missing values. The average decrease is greatest for the Apportionment degree A. For price-weighted versions $A_s(p_1^*)$ and $A_s(p_2^*)$ the decrease is approximatelly the same, but the values computed for $A_s(p_2^*)$ are at somewhat higher level. It is remarkable that although the percentage of generated missing values is relatively high, the Apportionment degree indicates quite good fit. For example, if the percentage of generated missing values is as high as 40 % the average value of the Apportionment index is approximatelly 0.89 with very narrow range of values (see Figure 2 in the Appendix). Thus even large

departures from the target table gave quite high values of the measure. This is not a very good property of a statistic, but it may make sense in practical applications where the values of the target table may not be possible to attain exactly. However, some kind of rough measure is needed to relate the target table to the observed one. Note that this measure compares only the relative values of the observed and target tables. Therefore large departures in the absolute values may not be noticed.



(a) Mean curves of A, $A_s(p_1^*)$ and $A_s(p_2^*)$. (b) Mean curves of p-values of χ^2, $\chi^2(p_1^*)$ and $\chi^2(p_2^*)$.

Figure 1. Mean curves of the simulation study.

Statistically the χ^2-statistic and its price-weighted versions performed better. When the percentage of missing values is 20 % or more, these statistics clearly reject the null hypothesis of the fit of the observed and target tables (Figure 1b). The performance of each of the χ^2-statistics follows approximately the similar pattern (see also Figure 3 in the Appendix).

The simulation in this section was carried out by using R computing environment (see e.g. http://www.r-project.org/).

4. CONCLUDING REMARKS

In this paper we study the use of the apportionment degree and the χ^2-statistic and their price-weighted versions when measuring the fit of the output and target tables. This comparison shows that the apportionment degree clearly measures the difference in relative values whereas the χ^2-statistic also observes the differences in absolute values. The idea of using prices as weights leads us to the use of the theory of index numbers for measuring the goodness of the bucking outcome which is a topic for future research.

5. ACKNOWLEDGMENTS

This work was financed by the Academy of Finland under the project entitled "Advanced Methods for Computer-Aided Bucking of Scots Pine". We are thankful to an anonymous referee for critically reading the manuscript and offering helpful suggestions which improved the presentation of our ideas.

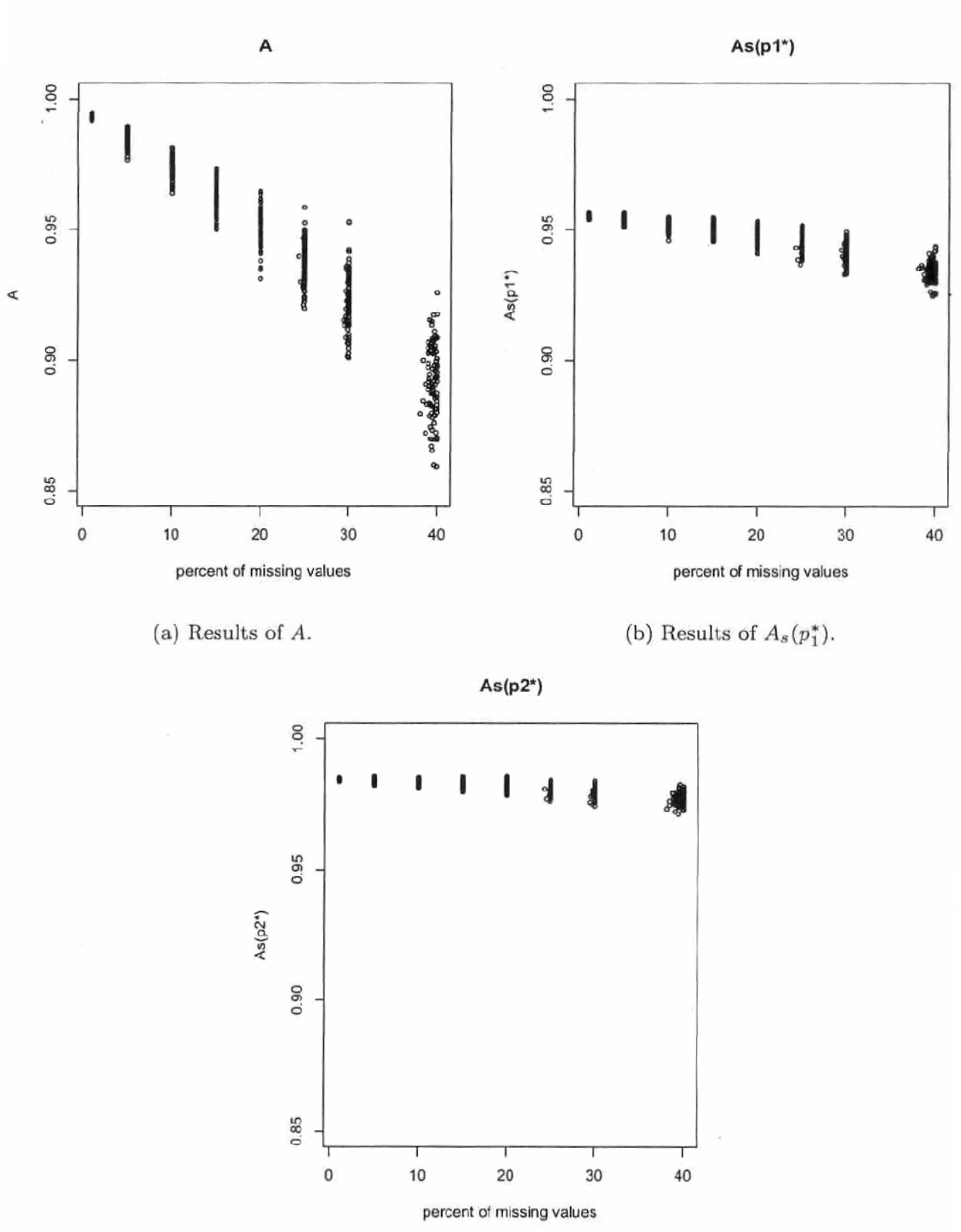
REFERENCES

LISKI, E.P. and T. NUMMI (1995): “Prediction of tree stems to improve efficiency in automatized harvesting of forests”, **Scandinavian Journal of Statistics**, 22, 255-269.

NÄSBERG, M. (1985): “Mathematical programming models for optimal log bucking”, **Linköping Studies in Science and Technology**, Dissertation No. 132.

RAO, C.R. (1973): **Linear Statistical Inference and its Applications**, Second Edition, Wiley, New York.

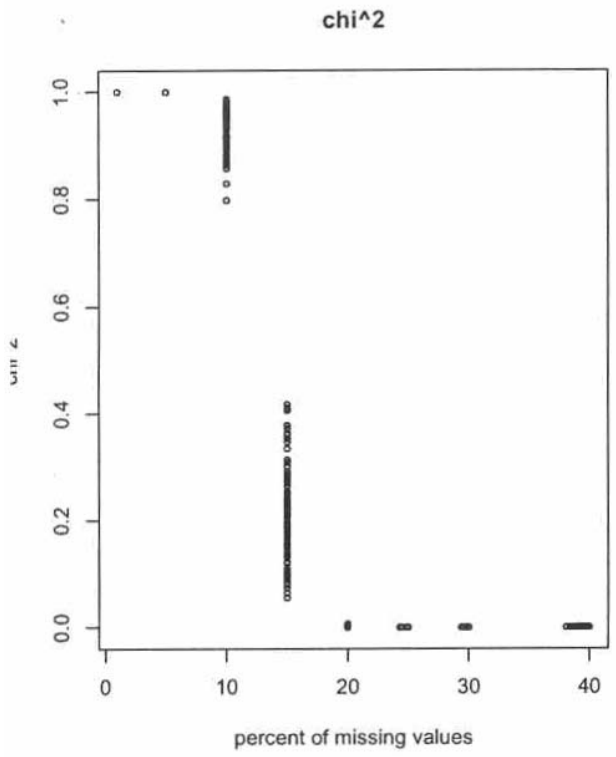
APPENDIX



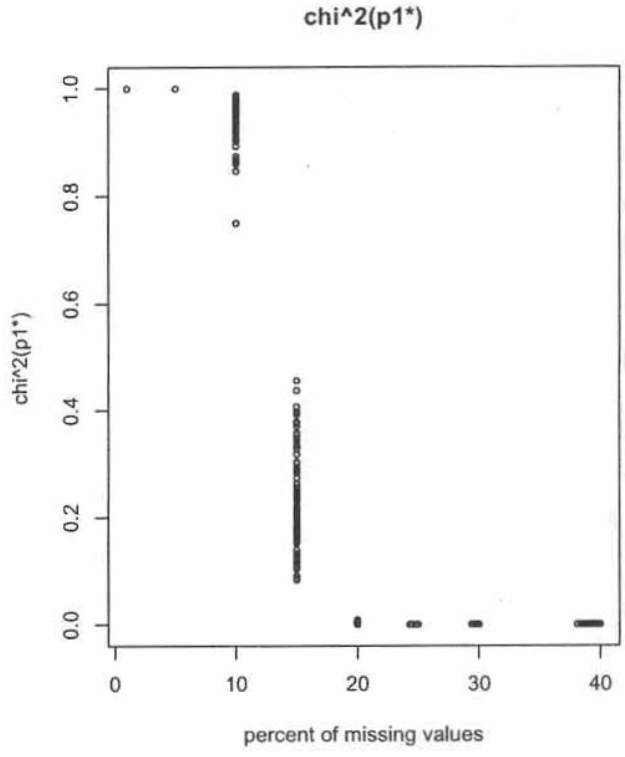
(a) Results of A.

(b) Results of $A_s(p_1^*)$.

Figure 2. Results of the simulation study for Apportionment degree and its scaled price-weighted versions.



(a) Associated p-values of χ^2.



(b) Associated p-values of $\chi^2(p_1^*)$.

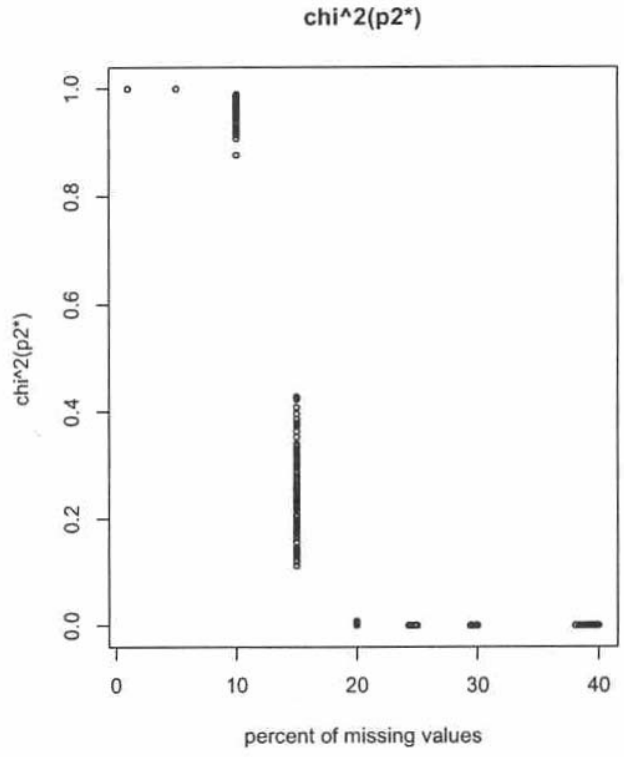


Figure 3. Results of the simulation study of ordinary χ^2-test statistic and its price-weighted versions.

ON A FAMILY OF APPORTIONMENT INDICES AND ITS LIMITING PROPERTIES

Revised version[1]

Bikas K. Sinha

Indian Statistical Institute, Kolkata, India

Laura Koskela

University of Tampere,Finland

Tapio Nummi

University of Tampere,Finland

In the harvesting technique prevailing in Scandinavia, tree stems are converted into smaller logs immediately at harvest. Modern sawmills attempt to operate according to customers' special needs rather than simply minimize production costs. Since the annual production of saw timber in Scandinavia is in the tens of millions of cubic meters, proper measurement of the goodness of the bucking outcome is of crucial importance. The outcome of the bucking operation can be considered as a multidimensional table of tree species, quality grades, prices and length and diameter classes. The prevailing method of measuring the outcome is the so-called Apportionment Index (AI), which is calculated from the relative proportions of the observed and target tables. Recently, some statistical properties of the AI have been studied and alternative means of measuring the bucking outcome have been suggested (Nummi et al., 2005 and Sinha et al., 2005). In this article we extend the definition of AI and examine its limiting properties.

Key Words: Apportionment degree, Price-weighted index, Standardization, Moment inequalities

1. INTRODUCTION

Modern sawmills seek to develop their production strategies based on customer demands in terms of the distribution of logs of various diameter-length specifications. The quality of the

[1]All results and computations presented in this paper were first published in the article: Sinha, B.K., Koskela, L., and Nummi. T. (2005). On a Family of Apportionment Indices and its Limiting Properties. *IAPQR transactions* **30**(2): 65-87.

actual harvesting operation has mainly been measured by calculating the relationship between the demand log distribution and the actual production distribution. A practice particularly commonly adopted in Scandinavia is to measure the bucking outcome by the so-called *Apportionment Degree*, or *Apportionment Index* (AI). This measure was introduced in forestry by the Swedish mathematician Bergstrand in the mid-1980s, when the first steps were taken in developing automatic bucking systems for forest harvesters. The main idea is to compare the relative proportions of the demand and target distributions (e.g. Bergstrand 1989). While there have been attempts to understand this index (Kivinen et al. 2005 and Nummi et al. 2005), a serious theoretical foundation is still lacking. Several alternatives to this measure are proposed in Kirkkala et al. (2000) and Malinen and Palander (2004). Only recently, statistical analysis of the AI has been initiated (Sinha et al. 2005).

Suppose that $\mathbf{X}^* = \left(x_{ij}^*\right)$ refers to the *proportional* output distribution and $\boldsymbol{\theta}^* = \left(\theta_{ij}^*\right)$ to the *proportional* demand distribution, i.e. both $\mathbf{X}^*$ and $\boldsymbol{\theta}^*$ refer to $m \times n$ matrices corresponding to diameter-length specifications and the elements satisfy the conditions $0 < x_{ij}^*, \theta_{ij}^* < 1 \ \ \forall\, i, j$ and $\sum_{i=1}^m \sum_{j=1}^n x_{ij}^* = \sum_{i=1}^m \sum_{j=1}^n \theta_{ij}^* = 1$. Then the AI is defined as

$$AI = 1 - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} |x_{ij}^* - \theta_{ij}^*|. \tag{1}$$

A little reflection shows that AI can also be written as

$$AI = \sum_{i=1}^{m} \sum_{j=1}^{n} \min(x_{ij}^*, \theta_{ij}^*), \tag{2}$$

since $\sum_{i=1}^m \sum_{j=1}^n x_{ij}^* = \sum_{i=1}^m \sum_{j=1}^n \theta_{ij}^* = 1$. The price-weighted version of AI is here denoted by AI^w and is defined as

$$AI^w = \sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij}^* \min(x_{ij}^*, \theta_{ij}^*), \tag{3}$$

where $\mathbf{P}^* = \left(p_{ij}^*\right)$ refers to the $m \times n$ matrix of *proportional* price distribution corresponding to diameter-length specifications.

We organize the rest of this paper as follows. In Section 2, we suggest generalizations of AI for both versions: unweighted and price-weighted. We also discuss appropriate standardizations for each of these measures. In Section 3, we first present some results related to moments and moment inequalities, then apply these inequalities to derive various properties of the standardized generalized measures of AI. A numerical illustration of the theoretical results is considered in Section 4. Section 5 discusses further issues in the standardization of the measures and presents additional pertinent results. In Section 6 we provide some concluding remarks and outline further work in this direction.

2. GENERALIZATION AND STANDARDIZATION OF THE MEASURES

2.1. On a Family of *AI*'s and its standarization

The traditional Apportionment Index (2) can be seen as a special case in the context of a whole family of AI's defined as

$$AI(k) = \sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n}\left[\min(x_{ij}^*, \theta_{ij}^*)\right]^k} = \sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n} {d_{ij}^*}^k}, \quad 0 < k < \infty, \tag{4}$$

where $d_{ij}^* = \min(x_{ij}^*, \theta_{ij}^*)$. It is easily seen that the traditional AI is simply a special case of $AI(k)$, i.e. $AI(1) = AI$.

Generally, $AI(k)$ cannot be interpreted on the same scale as the original Apportionment Index. To make the measures comparable, we must consider how to standardize $AI(k)$. The minimum value of AI is reached when all the logs fall into the diameter-length class of the smallest target proportion. This kind of scenario is referred to as a perfect mismatch. Clearly, $AI(k) \to 0 \ \ \forall\ k > 0$ as $d_{ij}^* \to 0 \ \ \forall\ i, j$, and this happens when $\min(\theta_{11}^*, \ldots, \theta_{mn}^*) \to 0$. The maximum index value is reached as the output and target matrices are equal (perfect match), i.e. $d_{ij}^* = \theta_{ij}^* \ \ \forall\ i, j$. It is then straightforward to show that in the case of a perfect match $AI(k) = \sqrt[k]{\sum\sum {\theta_{ij}^*}^k}, \ \ k > 0$, and more generally,

$$0 \le AI(k) \le \sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n} {\theta_{ij}^*}^k}, \quad k > 0. \tag{5}$$

The standardized generalized Apportionment Index can now be defined as

$$AI_s(k) = \frac{\sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n} {d_{ij}^*}^k}}{\sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n} {\theta_{ij}^*}^k}} = \frac{AI(k; d^*)}{AI(k; \theta^*)}, \quad k > 0. \tag{6}$$

The above notations $AI(k; d^*) = \sqrt[k]{\sum\sum {d_{ij}^*}^k}$ and $AI(k; \theta^*) = \sqrt[k]{\sum\sum {\theta_{ij}^*}^k}$ will be intensively used in the sequel.

2.2. On a Family of Price-weighted *AI*'s

In a similar manner, the price-weighted apportionment measure (3) can be generalized into a whole family of price-weighted Apportionment Indices defined as

$$AI^w(k) = \sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij}^* \left[\min(x_{ij}^*, \theta_{ij}^*)\right]^k}, \quad 0 < k < \infty. \tag{7}$$

Again, it is easily seen that $AI^w(1) = \sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij}^* d_{ij}^* = AI^w$. This time, using the same arguments as above, it turns out that

$$0 \leq AI^w(k) \leq \sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij}^* {\theta_{ij}^*}^k}, \quad k > 0, \tag{8}$$

and the standardized price-weighted generalized Apportionment Index can be defined as

$$AI_s^w(k) = \frac{\sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij}^* {d_{ij}^*}^k}}{\sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n} p_{ij}^* {\theta_{ij}^*}^k}} = \frac{AI^w(k;d^*)}{AI^w(k;\theta^*)}, \quad k > 0. \tag{9}$$

The notations $AI^w(k;d^*) = \sqrt[k]{\sum\sum p_{ij}^* {d_{ij}^*}^k}$ and $AI^w(k;\theta^*) = \sqrt[k]{\sum\sum p_{ij}^* {\theta_{ij}^*}^k}$ above will be intensively used in the sequel. Note that in a case of equal prices, i.e. $p_{ij}^* = \frac{1}{mn}\ \forall\ i,j$, the standardized versions of the traditional and the price-weighted indices are equal, i.e. $AI_s^w(k) = AI_s(k)$.

Illustrative examples for these computations are given in Section 4.

3. LIMITING PROPERTIES OF THE MEASURES

3.1. Moments of a discrete distribution and Liaponouv's inequality

Let Y be a positive-valued discrete random variable with finite moments having the distribution $[(y_i, \pi_i) \mid 1 \leq i \leq T]$, where π_i is the probability associated with the value y_i. The kth $(k > 0)$ moment of Y is

$$\gamma_k = E(Y^k) = \sum_{i=1}^{T} y_i^k \pi_i. \tag{10}$$

Recall that Liaponouv's inequality states that

$$\gamma_1 \leq \sqrt{\gamma_2} \leq \sqrt[3]{\gamma_3} \leq \sqrt[4]{\gamma_4} \leq \ldots \qquad \text{i.e.,} \qquad \gamma_k^{1/k} \uparrow \text{ in } k > 0. \tag{11}$$

Further, it is known that for a positive-valued discrete random variable Y with the above distribution

$$\sqrt[k]{\gamma_k} = \sqrt[k]{\sum_{i=1}^{T} y_i^k \pi_i} = \begin{cases} \min(y_i) & \text{as } k \to -\infty \\ HM & \text{as } k = -1 \\ GM & \text{as } k \to 0 \\ AM & \text{as } k = 1 \\ \max(y_i) & \text{as } k \to \infty \end{cases}, \tag{12}$$

where HM, GM and AM refer to the harmonic mean, geometric mean and arithmetic mean of y-values, respectively.

3.2. Relation of the Apportioment Index to moments of a discrete distribution

The above aspects of moments and moment inequalities will be useful in our subsequent analysis. To see this, we first note that $AI(k;d^*) = \sqrt[k]{\sum\sum {d_{ij}^*}^k}$ bears a relation to γ_k for discrete distribution if we can specify a probability distribution. By taking the probablities as $\pi_{ij} = \frac{1}{mn}$, we may write

$$\gamma_k(d^*) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} {d_{ij}^*}^k}{mn}. \tag{13}$$

We here use the notation $\gamma_k(d^*)$ for the kth moment to emphasize that it is computed for the d_{ij}^*-values. By Liaponouv's inequality we now get

$$\sqrt[k]{\gamma_k(d^*)} = \sqrt[k]{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n} {d_{ij}^*}^k}{mn}} \uparrow \text{ in } k>0 \quad \text{i.e.,} \quad \frac{AI(k;d^*)}{\sqrt[k]{mn}} \uparrow \text{ in } k>0. \tag{14}$$

In the case of the price-weighted index $AI^w(k;d^*)$, the proportional price distribution may very well serve as a probability distribution, since $0 < p_{ij}^* < 1$ for all i,j and $\sum\sum p_{ij}^* = 1$. Writing the generalized price-weighted measure as $AI^w(k;d^*) = \sqrt[k]{\sum\sum p_{ij}^* {d_{ij}^*}^k} = \sqrt[k]{\gamma_k(d^*;p^*)}$, Liaponouv's inequality gives

$$AI^w(k;d^*) = \sqrt[k]{\gamma_k(d^*;p^*)} \uparrow \text{ in } k>0. \tag{15}$$

The notation $\gamma_k(d^*;p^*)$ is used here to emphasize that the kth moment is computed for the d_{ij}^*-values taking prices as the corresponding probabilities. The above result also holds for $AI^w(k;\theta^*) = \sqrt[k]{\sum\sum p_{ij}^* {\theta_{ij}^*}^k} = \sqrt[k]{\gamma_k(\theta^*;p^*)}$.

3.3. On the properties of the non-weighted measures

We first refer to the standardized measure in (6) and rewrite it as

$$AI_s(k) = \frac{AI(k;d^*)}{AI(k;\theta^*)} = \frac{\sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n} {d_{ij}^*}^k}}{\sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n} {\theta_{ij}^*}^k}} = \frac{\sqrt[k]{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n} {d_{ij}^*}^k}{mn}}}{\sqrt[k]{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n} {\theta_{ij}^*}^k}{mn}}} = \frac{\sqrt[k]{\gamma_k(d^*)}}{\sqrt[k]{\gamma_k(\theta^*)}} = \sqrt[k]{\frac{\gamma_k(d^*)}{\gamma_k(\theta^*)}}. \tag{16}$$

Although $AI_s(k;d^*)$ is originally defined only for $k>0$, the relation to the moments shows that we *may* extend the range of definition $AI_s(k)$ to cover the entire real line, namely $-\infty < k < \infty$. However, the extension may *not* have any statistical interpretation for $k<0$. In view of (12), we may deduce that

$$AI_s(k) = \frac{\sqrt[k]{\gamma_k(d^*)}}{\sqrt[k]{\gamma_k(\theta^*)}} \to \begin{cases} \frac{d_{min}^*}{\theta_{min}^*} & \text{as } k \to -\infty \\ \frac{HM(d^*)}{HM(\theta^*)} & \text{as } k=-1 \\ \frac{GM(d^*)}{GM(\theta^*)} & \text{as } k \to 0 \\ \frac{AM(d^*)}{AM(\theta^*)} = mn\bar{d}^* & \text{as } k=1 \\ \frac{d_{max}^*}{\theta_{max}^*} & \text{as } k \to \infty \end{cases}. \tag{17}$$

In the above, we have used the abbreviations $d^*_{min} = \min(d^*_{11}, \ldots, d^*_{mn})$, $d^*_{max} = \max(d^*_{11}, \ldots, d^*_{mn})$, $HM(d^*) = \frac{mn}{\sum\sum \frac{1}{d^*_{ij}}}$, $GM(d^*) = \sqrt[mn]{\prod\prod d^*_{ij}}$ and $AM(d^*) = \frac{\sum\sum d^*_{ij}}{mn} = \bar{d}^*$. Corresponding notations are used for θ^*.

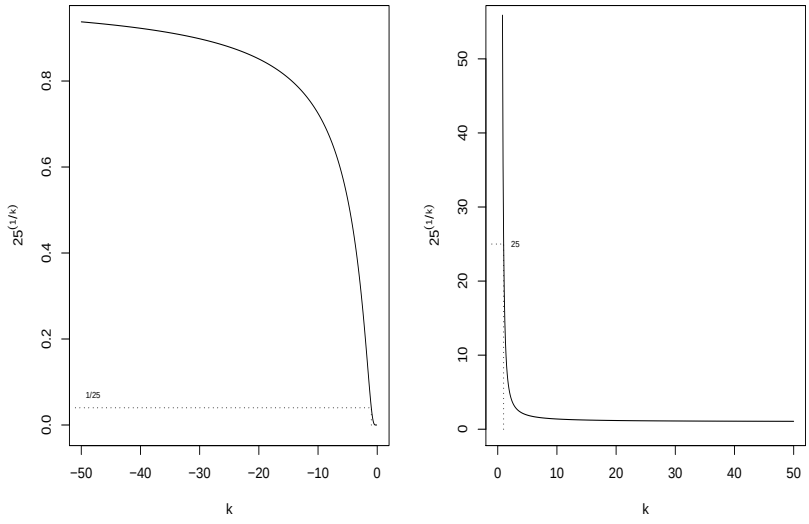
The traditional index can be expressed as

$$AI(k; d^*) = (mn)^{\frac{1}{k}} \sqrt[k]{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n} {d^*_{ij}}^k}{mn}} = (mn)^{\frac{1}{k}} \sqrt[k]{\gamma_k(d^*)}, \quad k > 0. \tag{18}$$

Before any extension to the range of $AI(k; d^*)$ can be made, the nature of the function $(mn)^{\frac{1}{k}}$ for $k < 0$ needs to be studied. A closer look at the function shows that

$$\begin{aligned} &(i)\ (mn)^{\frac{1}{k}} \ \downarrow \text{ from 1 to } \frac{1}{mn} \text{ for } k \le -1, \\ &(ii)\ (mn)^{\frac{1}{k}} \ \downarrow \text{ from } \frac{1}{mn} \text{ to 0 for } -1 \le k < 0 \text{ and} \\ &(iii)\ (mn)^{\frac{1}{k}} \ \downarrow \text{ from } \infty \text{ to 1 for } 0 < k < \infty. \end{aligned} \tag{19}$$

Figure 1 illustrates the behaviour of $(mn)^{\frac{1}{k}}$ for $m = n = 5$.



(a) $(mn)^{\frac{1}{k}}$ for $k < 0$ (b) $(mn)^{\frac{1}{k}}$ for $k > 0$

FIG. 1. $(mn)^{\frac{1}{k}}$ vs. k with $m = n = 5$

Although the quantity $\sqrt[k]{\gamma_k(d^*)}$ is well-defined for $-\infty < k < \infty$, $AI(k; d^*)$ has a discontinuity point at $k = 0$. It is, however, a piecewise decreasing function in both $(-\infty, 0)$ and

$(0, \infty)$ and it follows from (19) that

$$\tfrac{1}{mn}\sqrt[k]{\gamma_k(d^*)} \leq AI(k;d^*) \leq \sqrt[k]{\gamma_k(d^*)} \text{ for } k \leq -1$$

and

$$0 \leq AI(k;d^*) \leq \frac{1}{mn}\sqrt[k]{\gamma_k(d^*)} \text{ for } -1 \leq k < 0.$$

An extension of the range of $AI(k;d^*)$ for $-\infty < k < \infty$ is hence possible also for $AI(k;d^*)$. We may now deduce that

$$AI(k;d^*) \rightarrow \begin{cases} d^*_{min} & \text{as } k \rightarrow -\infty \\ \frac{HM(d^*)}{mn} & \text{as } k = -1 \\ 0 & \text{as } k \rightarrow 0- \\ \infty & \text{as } k \rightarrow 0+ \\ mnAM(d^*) & \text{as } k = 1 \\ d^*_{max} & \text{as } k \rightarrow \infty \end{cases} . \tag{20}$$

Analogous limiting expressions can also be written for $AI(k;\theta^*)$. In this context, we refer to Table 5 and Figure 2 in Section 4.

Remark 1 It is interesting to note that whereas both $AI(k;d^*)$ and $AI(k;\theta^*)$ are discontinuous at $k = 0$, the standardized measure $AI_s(k)$ is continuous everywhere, including $k = 0$. It is readily seen that $AI_s(k) \rightarrow \frac{GM(d^*)}{GM(\theta^*)}$ as $k \rightarrow 0+$. Further, it can be argued that the same limit prevails as $k \rightarrow 0-$.

3.4. On the properties of the price-weighted measures

The standardized price-weighted measure (9) can be rewritten as

$$AI^w_s(k) = \frac{AI^w(k;d^*)}{AI^w(k;\theta^*)} = \frac{\sqrt[k]{\gamma_k(d^*;p^*)}}{\sqrt[k]{\gamma_k(\theta^*;p^*)}} = \sqrt[k]{\frac{\gamma_k(d^*;p^*)}{\gamma_k(\theta^*;p^*)}}. \tag{21}$$

Hence, in view of (12)

$$AI^w_s(k) = \sqrt[k]{\frac{\gamma_k(d^*;p^*)}{\gamma_k(\theta^*;p^*)}} \rightarrow \begin{cases} \frac{d^*_{min}}{\theta^*_{min}} & \text{as } k \rightarrow -\infty \\ \frac{HM(d^*;p^*)}{HM(\theta^*;p^*)} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{p^*_{ij}}{\theta^*_{ij}}}{\sum_{i=1}^{m}\sum_{j=1}^{n}\frac{p^*_{ij}}{d^*_{ij}}} & \text{as } k = -1 \\ \frac{GM(d^*;p^*)}{GM(\theta^*;p^*)} = \frac{\prod_{i=1}^{m}\prod_{j=1}^{n} {d^*_{ij}}^{p^*_{ij}}}{\prod_{i=1}^{m}\prod_{j=1}^{n} {\theta^*_{ij}}^{p^*_{ij}}} & \text{as } k \rightarrow 0 \\ \frac{AM(d^*;p^*)}{AM(\theta^*;p^*)} & \text{as } k = 1 \\ \frac{d^*_{max}}{\theta^*_{max}} & \text{as } k \rightarrow \infty \end{cases} , \tag{22}$$

where $HM(d^*;p^*)$, $GM(d^*;p^*)$ and $AM(d^*;p^*)$ refer to the price-weighted harmonic mean, geometric mean and arithmetic mean, respectively. These means are computed for the values $d^*_{ij} = \min(x^*_{ij}, \theta^*_{ij})$ taking prices as the corresponding probabilities. Similarly, $HM(\theta^*;p^*)$, $GM(\theta^*;p^*)$ and $AM(\theta^*;p^*)$ refer to the price-weighted means computed for the target values.

4. ILLUSTRATIVE EXAMPLES

4.1. Illustrations on the behaviour of the non-weighted measures

We now illustrate the behaviour of the non-weighted measures as a function of k for fixed target and output matrices. The target matrix of Nummi et al. (2005) is used here and is given in Table 1. The output matrix is obtained from the target matrix by randomly dropping 10% of logs in the target matrix (cf. Nummi et al. 2005) and is displayed in Table 2. We only show the first four decimals, which explain the property of proportions not adding exactly up to one. The matrix $\mathbf{D}^* = \left(d^*_{ij}\right)$, where $d^*_{ij} = \min(x^*_{ij}, \theta^*_{ij})$, is shown in Table 3.

TABLE 1.

Target matrix $\boldsymbol{\theta}^*$.

0.028	0.016	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

TABLE 2.

Output matrix $\mathbf{X}^*$.

0.0289	0.0144	0.0622	0.0411	0.0456
0.0378	0.0067	0.0689	0.0467	0.0356
0.0167	0.0456	0.0344	0.0456	0.0589
0.0222	0.0400	0.0400	0.0444	0.0633
0.0178	0.0267	0.0478	0.0544	0.0544

TABLE 3.

Matrix $\mathbf{D}^*$.

0.0280	0.0144	0.0580	0.0411	0.0450
0.0370	0.0067	0.0650	0.0450	0.0356
0.0167	0.0456	0.0344	0.0440	0.0550
0.0220	0.0390	0.0390	0.0440	0.0590
0.0178	0.0267	0.0470	0.0540	0.0520

To study the behaviour of $AI(k;d^*)$, $AI(k;\theta^*)$ and $AI_s(k)$, it is necessary to make some computations based on the target and $\mathbf{D}^*$ matrices. These we show in Table 4. The computations of the limits of the apportionment measures in (17) and (20) are reported in Table 5. Figures 2 and 3 are provided to illustrate the behaviour of $AI(k;d^*)$ respective $AI_s(k)$ for some wide

range of k. Figure 2 clearly shows that $AI(k;d^*)$ has a discontinuity point at $k=0$ and that it approaches d^*_{min} and d^*_{max} as k tends to $-\infty$ and ∞, respectively. Figure 3a shows, however, that $AI_s(k)$ forms a nonsymmetric S-shaped curve as k runs through the whole range. Figure 3b takes a closer look at the measure at the interval $k \in [-1,1]$ showing that $AI_s(k)$ forms a continuous curve with the limit $GM(d^*)/GM(\theta^*)$ as $k \to 0$ (cf. Table 5).

TABLE 4.
Summary statistics computed from the target and output matrices.

Statistics	d^*	θ^*
min	0.0067	0.0160
HM	0.0293	0.0338
GM	0.0349	0.0371
AM	0.0389	0.0400
max	0.0650	0.0650

TABLE 5.
Behaviour of the generalized and standardized non-weighted measures as k varies.

k	$AI(k;d^*)$	$AI(k;\theta^*)$	$AI_s(k)$
$-\infty$	0.0067	0.0160	0.4167
-1	0.012	0.0014	0.8667
0	–	–	0.9411
1	0.9719	1.0000	0.9719
∞	0.0650	0.0650	1.0000

Remark 2 $\frac{d^*_{max}}{\theta^*_{max}} \leq 1$, and the equality holds if and only if $x^*_{ij_{max}} \geq \theta^*_{ij_{max}}$, where ij_{max} denotes the cell index of the largest target value.

Remark 3 It is observed in our computation that

$$\frac{d^*_{min}}{\theta^*_{min}} < \frac{HM(d^*)}{HM(\theta^*)} < \frac{GM(d^*)}{GM(\theta^*)} < \frac{AM(d^*)}{AM(\theta^*)} < \frac{d^*_{max}}{\theta^*_{max}} = 1,$$

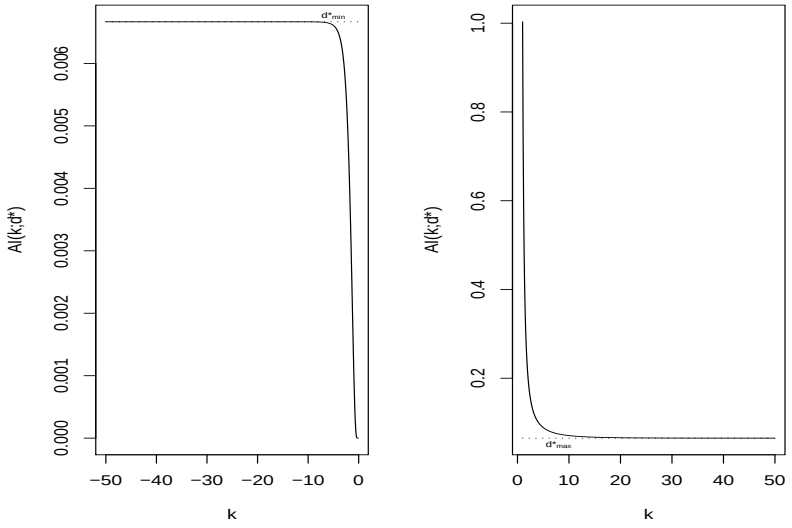
though, in general terms, it is easily seen that each of the ratios is ≤ 1.

4.2. Illustrations of the behaviour of the price-weighted measures

To illustrate the behaviour of $AI^w_s(k;d^*)$ for different values of k, we use the same target and output matrices as above and the price matrix presented in Nummi et al. (2005). The price matrix is displayed in Table 6.

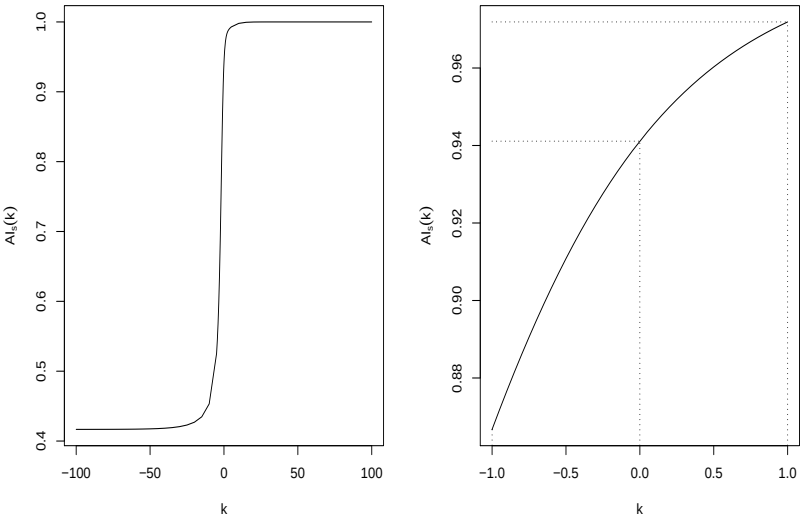
TABLE 6.
Price matrix $\mathbf{P}^*$.

0.0279	0.0287	0.0292	0.0301	0.0304
0.0345	0.0356	0.0362	0.0373	0.0376
0.0401	0.0412	0.0421	0.0434	0.0437
0.0434	0.0448	0.0457	0.0468	0.0473
0.0446	0.0460	0.0468	0.0482	0.0485



(a) $AI(k; d^*)$ for $k < 0$. (b) $AI(k; d^*)$ for $k > 0$.

FIG. 2. $AI(k; d^*)$ vs. k for a given target matrix $\boldsymbol{\theta}^*$ and output matrix $\mathbf{X}^*$.



(a) $AI_s(k)$ vs. $k \in [-100, 100]$. (b) $AI_s(k)$ for $k \in [-1, 1]$.

FIG. 3. $AI_s(k)$ vs. k.

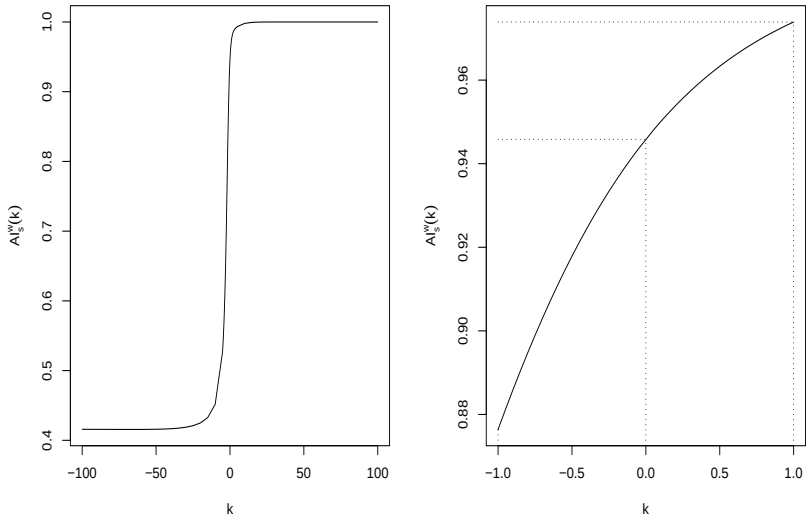
The computations for the minimum, maximum and price-weighted means for the matrices $\mathbf{d}^*$ and $\boldsymbol{\theta}^*$ are shown in Table 7. To examine the behaviour of $AI_s^w(k)$ we compute the limits given in (22). These computations are reported in Table 8. We also provide Figure 4 to further illustrate the behaviour of $AI_s^w(k)$ as k runs through some wide range of values. As was the case with $AI_s(k)$, also $AI_s^w(k)$ forms a continuous nonsymmetric S-shaped curve. According to the derivations in (22), $AI_s^w(k) \to GM(d^*;p^*)/GM(\theta^*;p^*)$ as $k \to 0$. Figure 4 and the computations in Table 8 verify this.

TABLE 7.
Summary statistics computed from the target and output matrices with respect to price.

Statistics	$d^*;p^*$	$\theta^*;p^*$
min	0.0067	0.0160
HM	0.0302	0.0344
GM	0.0356	0.0376
AM	0.0394	0.0404
max	0.0650	0.0650

TABLE 8.
Behaviour of the standardized price-weighted measure as k varies.

k	$AI_s^w(k)$
$-\infty$	0.4167
-1	0.8764
0	0.9458
1	0.9739
∞	1.0000



(a) $AI_s^w(k)$ for $k \in [-100, 100]$.

(b) $AI_s^w(k)$ for $k \in [-1, 1]$.

FIG. 4. $AI_s^w(k)$ vs. k.

5. FURTHER STUDIES ON STANDARDIZATION OF THE MEASURES

5.1. On standardization of the non-weighted measure

So far in our study, the attainable minimum of $AI(k;d^*)$ is taken to be independent of the given target matrix. The independence follows from assuming that $\min(\theta^*_{11},\ldots,\theta_{mn}) \to 0$, i.e. that the smallest target value is indeed almost zero. Since the minimum corresponds to a perfect mismatch, i.e. all the logs fall into the diameter-length class with the smallest target value, it is easily seen that $AI(k;d^*) \to 0$ only under the above assumption. Another approach to standardize the measure would hence be to adjust the minimun of $AI(k;d^*)$ on the *given* target matrix. Suppose that $\theta^*_{ij_{min}}$ represents the smallest (positive) element in the target matrix $\boldsymbol{\theta}^*$. Under a perfect mismatch $AI(k;d^*)$ then reduces to

$$AI(k;d^*) = \sqrt[k]{\min(\theta^*_{ij})^k} = \theta^*_{ij_{min}}.$$

Adjusting the minimun of $AI(k;d^*)$ on the *given* target matrix, the "adjusted" standardized non-weighted measure, $AI_s(k)_{adj.}$, then becomes

$$AI_s(k)_{adj.} = \frac{AI(k;d^*) - \theta^*_{ij_{min}}}{AI(k;\theta^*_{ij}) - \theta^*_{ij_{min}}}, \quad k > 0.$$

Defining $d^{**}_{ij} = \frac{d^*_{ij}}{\theta^*_{ij_{min}}}$ and $\theta^{**}_{ij} = \frac{\theta^*_{ij}}{\theta^*_{ij_{min}}}$ for all i,j, we may rewrite $AI_s(k)_{adj.}$ as

$$AI_s(k)_{adj.} = \frac{\sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n} {d^*_{ij}}^k} - \theta^*_{ij_{min}}}{\sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n} {\theta^*_{ij}}^k} - \theta^*_{ij_{min}}} = \frac{\sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n} {d^{**}_{ij}}^k} - 1}{\sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n} {\theta^{**}_{ij}}^k} - 1} = \frac{\sqrt[k]{\gamma_k(d^{**})} - (mn)^{-\frac{1}{k}}}{\sqrt[k]{\gamma_k(\theta^{**})} - (mn)^{-\frac{1}{k}}}. \tag{23}$$

Recalling the behaviour of $(mn)^{\frac{1}{k}}$ in (19), it is easy to see that $(mn)^{-\frac{1}{k}} \uparrow$ from 0 to 1 for $0 < k < \infty$. Then,

$$\lim AI_s(k)_{adj.} \to \begin{cases} \frac{GM(d^{**})}{GM(\theta^{**})} = \frac{GM(d^*)}{GM(\theta^*)} & \text{as } k \to 0+ \\ \frac{mn\bar{d}^{**}-1}{mn\bar{\theta}^{**}-1} = \frac{mn\bar{d}^{*}-\theta^*_{ij_{min}}}{mn\bar{\theta}^{*}-\theta^*_{ij_{min}}} = \frac{\bar{d}^{*}-\frac{\theta^*_{ij_{min}}}{mn}}{\bar{\theta}^{*}-\frac{\theta^*_{ij_{min}}}{mn}} & \text{as } k \to 1 \\ \frac{d^{**}_{max}-1}{\theta^{**}_{max}-1} = \frac{d^*_{max}-\theta^*_{ij_{min}}}{\theta^*_{max}-\theta^*_{ij_{min}}} & \text{as } k \to \infty \end{cases}. \tag{24}$$

Notations $\bar{d}^*$, $\bar{d}^{**}$, $\bar{\theta}^*$ and $\bar{\theta}^{**}$ correspond to arithmetic means.

In Figure 5 we compare the two standardized measures $AI_s(k)$ and $AI_s(k)_{adj.}$ for $k > 0$ and the given target and output matrices. The figure shows that the measures are only slightly different. Our computations support this conclusion, since the limiting values of $AI_s(k)$ and $AI_s(k)_{adj.}$ are the same as $k \to 0+$ and $k \to \infty$ (see the values of $AI_s(k)$ in Table 5). However, as k tends to one, $AI_s(k)_{adj.} \to 0.9711$ and $AI_s(k) \to 0.9719$.

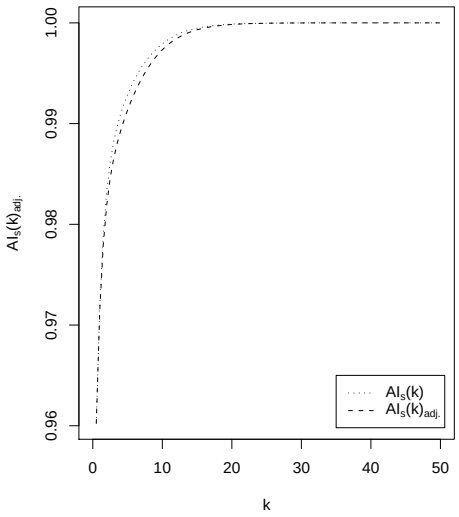


FIG. 5. $AI_s(k)$ and $AI_s(k)_{adj.}$ vs. k

Remark 4 As noted in Remark 2, when $x^*_{ij_{max}} \geq \theta^*_{ij_{max}}$, then $d^*_{max} = \theta^*_{max}$ and hence $AI_s(k)$ and $AI_s(k)_{adj.}$ are equal to 1 as $k \to \infty$.

5.2. On standardization of the price-weighted measure

Assuming that all target values are positive, $AI^w(k; d^*)$ achieves its minimum as $x^*_{ij} \to 1$ for the smallest $z_{ij} = p^*_{ij}\theta^*_{ij}$ and $x^*_{ij} \to 0$ for all other values. The value of $AI^w(k; d^*)$ then equals the smallest z_{ij}, which will be denoted in the sequel by $z_{ij_{min}}$. Following the derivations in Section 5.1, we now define the "adjusted" price-weighted standardized measure, which can be written as

$$AI^w_s(k)_{adj.} = \frac{AI^w(k; d^*) - z_{ij_{min}}}{AI^w(k; \theta^*_{ij}) - z_{ij_{min}}} \quad k > 0.$$

By rewriting $AI^w_s(k)_{adj.}$ as

$$AI^w_s(k)_{adj.} = \frac{\sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n} p^*_{ij} {d^*_{ij}}^k} - z_{ij_{min}}}{\sqrt[k]{\sum_{i=1}^{m}\sum_{j=1}^{n} p^*_{ij} {\theta^*_{ij}}^k} - z_{ij_{min}}} = \frac{\sqrt[k]{\gamma_k(d^*; p^*)} - z_{ij_{min}}}{\sqrt[k]{\gamma_k(\theta^*; p^*)} - z_{ij_{min}}}, \tag{25}$$

we may deduce that

$$\lim AI_s^w(k)_{adj.} \rightarrow \begin{cases} \frac{GM(d^*;p^*)-z_{ij_{min}}}{GM(\theta^*;p^*)-z_{ij_{min}}} & \text{as } k \rightarrow 0+ \\ \frac{AM(d^*;p^*)-z_{ij_{min}}}{AM(\theta^*;p^*)-z_{ij_{min}}} & \text{as } k \rightarrow 1 \\ \frac{d^*_{max}-z_{ij_{min}}}{\theta^*_{max}-z_{ij_{min}}} & \text{as } k \rightarrow \infty \end{cases} . \quad (26)$$

Overall, little difference is seen between the two standardized measures $AI_s^w(k)$ and $AI_s^w(k)_{adj.}$ for $k > 0$ and the given matrices. The computation of the limits of $AI_s^w(k)_{adj.}$ in (26) give 0.9452, 0.9736 and 1.000 for $k \rightarrow 0+,\ 1$ and ∞, respectively. A comparison of these values to those of $AI_s^w(k)$ in Table 8 reveal, however, that the measures differ slightly even for the case $k \rightarrow 0+$.

Remark 5 Also for the price-weighted measure $d^*_{max} = \theta^*_{max}$, when $x^*_{ij_{max}} \geq \theta^*_{ij_{max}}$. In such a situation both $AI_s^w(k)$ and $AI_s^w(k)_{adj.}$ are equal to 1 as $k \rightarrow \infty$.

6. CONCLUSION

Following the concept and definition of the Apportionment Index, as suggested by Bergstrand (1990), we have here undertaken a study of the extension of AI and an examination of its limiting behaviour. This is a sequel to the recent studies reported in papers by Nummi et al. (2005) and Sinha et al. (2005) on this subject. As is pointed out in Nummi et al. (2005), it may be remarked that consideration of the cost of logs of different length-diameter combinations and of varying qualities is essential in such studies. This calls for a modification to the definition of AI and its generalizations. Results on the price-weighted version of the measure and its extensions are given in this paper.

Several alternatives to the Apportionment Index are proposed in papers by Kirkkala et al. (2000) and Malinen and Palander (2004). However, since the scaling of these alternative measures differs from that of the original AI, the standardization of these measures needs to be studied before the results can be compared e.g. to the traditional measure. The standardization and generalization of the measures as well as the price-weighted considerations of the proposed measures remain a topic for future research.

The introduction of the family of AI's may raise the question of the *optimal* choice of k for a given target matrix $\boldsymbol{\theta}^*$. Study in this direction is also left as a topic for future research.

7. ACKNOWLEDGEMENT

The research visit of Bikas K Sinha was made possible by the support of the Statistics Unit of the Department of Mathematics, Statistics and Philosophy in the University of Tampere as also by the Academy of Finland under the project entitled "Forest-level bucking optimization including transportation cost, product demand and stand characteristics", (Project 1104405).

REFERENCES

Bergstrand, K.-G. (1989). Fördelningsaptering med näroptimalmetoden – Reviderad version [Bucking to order with a close-to-optimal method – Revised version]. Forskningsstiftelsen Skogsarbeten. 1989-12-11. 11 p. [In Swedish.]

Kirkkala, A., Sikanen, L., Harstela, P., Ruha, T, and Tarnanen, T. (2000). Sakkosegmentoitu tavoitejakauma apteeraustuloksen arvioinnissa ja jakauma-asteen laskennassa [Assessing the goodness of the bucking outcome by penalty-segmented log demand distribution]. *Metsätieteen aikakauskirja 1/2000*: 59-61. [In Finnish.]

Kivinen, V.-P., Uusitalo, J., and 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.

Malinen, J., and Palander T. (2004). Metrics for distribution similarity applied to the bucking to demand procedure. *International Journal of Forest Engineering* **15**(1): 33-40.

Nummi, T., Sinha, B.K., and Koskela, L. (2005). Statistical Properties of the Apportionment Degree and Alternative Measures in Bucking Outcome. *Revista Investigación Operacional* **26**(3): 259–267.

Sinha, B.K., Koskela, L., and Nummi, T. (2005). On Some Statistical Properties of the Apportionment Index. *Revista Investigación Operacional* **26**(2): 169-179.

ON SOME STATISTICAL PROPERTIES OF THE APPORTIONMENT INDEX

Bikas K. Sinha[1], Indian Statistical Institute, India
Laura Koskela[2] and Tapio Nummi[3], University of Tampere, Finland

ABSTRACT
The purpose of this paper is to initiate a statistical analysis of the *Apportionment index*, a measure used in the context of forest harvesting to evaluate the fit between demand and supply distribution of logs. There have been some attempts to understand this index, but a serious theoretical foundation is still lacking. We briefly review the available literature and then proceed to investigate the properties of the index from a distributional point of view. This is mainly an exploratory article and we focus only on the cases of two and three log classes, i.e., locations. In the case of two locations we use the beta distribution for the random relative output variables; with three locations the random relative outputs are assumed to follow a singular Dirichlet distribution. Using this formulation it is possible to understand the statistical properties of the apportionment index.

Key words: Mean deviation, beta distribution, Dirichlet distribution, exchangeable distributions, median.

RESUMEN
El propósito de este trabajo es comenzar el análisis estadístico del índice del prorrateo, esta es una medida usada en el contexto de la tala en bosque para evaluar el ajuste entre la demanda y distribución del suministro de leña. Ha habido algunos esfuerzos por entender este índice, pero se necesita una base teórica seria. Discutimos brevemente la literatura existente y procedemos a investigar las propiedades del índice desde un punto de vista distribucional. Éste es fundamentalmente un artículo exploratorio y sólo enfocamos los casos de dos y tres clases de leña, es decir, locaciones. En el caso de dos clases usamos la distribución beta para las variables aleatorias relativas a la salida (output) del rendimiento; en tres locaciones al azar se asume que los rendimientos relativos siguen la distribución de Dirichlet singular. Usando esta formulación es posible entender las propiedades estadísticas del índice del prorrateo.

MSC: 62P12.

1. INTRODUCTION

Modern sawmills attempt to develop their production strategies based on customer demands in terms of the distribution of logs of various diameter - length specifications. The quality of the actual harvesting operation has mainly been measured by calculating the relationship between the demand log distribution and the actual production distribution. A very commonly adopted practice in Scandinavia is to measure the bucking outcome by the so-called *apportionment degree*, or *apportionment index* (AI). This measure was developed by the Swedish mathematician Bergstrand in the mid 1980s, when the first steps were taken in developing automatic bucking systems for forest harvesters. The main idea is to compare the relative proportions of the demand and target distributions (e.g. Bergstrand 1990). While there have been attempts to understand this index (Kivinen **et al**. 2003 and Nummi **et al**. 2004), a serious theoretical foundation is still lacking. Several extensions of this measure are proposed in Kirkkala et al. (2000) and Malinen & Palander (2004).

Our interest here is in the statistical analysis of the apportionment index AI. This kind of analysis may have many potential applications in harvesting. For example, in harvesting planning we may have many possible output distributions (stands) and we should be able to select the optimal one for a given target. Further, we may have many possible targets and we wish to know which is optimal for a given output. A proper understanding of the statistical properties of this measure is thus of great importance.

In Section 2 we initiate a statistical study of the AI based on the joint distribution of the random component outputs in the output matrix. In section 2.1 we take the analysis of the two log classes as a starting-point. in Section 2.2 we extend the analysis to the case of three locations and finally in section 3 some observations are made on the future course of action.

E-mail: [1]bksinha@isical.ac.in
[2]laura.koskela@uta.fi*
[3]tan@uta.fi*

Before closing the section, we record the general definition of AI. Suppose X refers to the *proportional* output distribution and θ refers to the *proportional* demand distribution. Then the AI is defined as

$$AI = 1 - \frac{1}{2}\sum_i \sum_j | X_{ij} - \theta_{ij} | . \quad (1)$$

After some simple manipulations we can show that AI can also be written as

$$AI = \sum_i \sum_j \min\{X_{ij} . \theta_{ij}\}. \quad (2)$$

since $\sum_i \sum_j X_{ij} = \sum_i \sum_j \theta_{ij} = 1$ and $\min(a, b) = \frac{a+b}{2} - \frac{|a-b|}{2}$ for two real numbers a and b.

2. JOINT DISTRIBUTION OF RANDOM OUTPUTS AND STATISTICAL ANALYSIS OF AI

2.1. The case of two locations

Let us assume for simplicity that there are only two locations, labeled “L1" and “L2", and their relative (i.e., proportional) demands are θ and 1 - θ, respectively. Let the relative random output generated at location L1 be denoted by X so that in location L2 the output generated is 1 - X. By the definition the AI is given by the formula (1) and in the case of two locations it can be written as

$$AI = 1 - | X - \theta |. \quad (3)$$

Because AI is now a random quantity we may look at the expected value of the AI given by the formula

$$E(AI) = 1 - E [| X - \theta |] . \quad (4)$$

At this stage we note that the relative random output X in L1 is a random variable defined over [0,1] and this explains the random nature of AI in (3). Our purpose is to maximize AI in some sense, since this will suggest maximum apportionment. By reason of the stochastic nature of AI, one possibility would be to attain a heavy right tail distribution for AI so that it will tend to be probabilistically large. In this paper, we use the notion of maximization in the averaged sense, i.e., we aim at maximizing the expectation of AI in (4). This is equivalent to minimizing E[|X-θ|]. Since the mean deviation is least when it is taken about the median, our goal is to recommend a distribution for X for which the median is the *known* target value of θ, say θ_0. Since X is distributed over [0,1], it is natural to express its distribution as a member of the family of beta distributions (B(x;α,β)) introduced below by the density in (5).

$$f(x; \alpha, \beta) = x^{\alpha-1}(1 - x)^{\beta-1} / B(\alpha, \beta); 0 < x < 1, \quad (5)$$

where B(α, β) is the beta integral defined for α > 0, β > 0. On this see e.g. Johnson **et al**. (1995), 2nd Edition, pp. 210-211; Kotz and Johnson (1982), Vol.1, pp. 228-229. We may thus seek to use the beta distribution for X with proper choice of the parameters α and β, determined by the condition that θ_0 serves as the median of the X distribution. In effect, we seek a solution for α and β, so as to satisfy

$$0.5 = \int_0^{\theta_0} f(x; \alpha,\beta)\, dx . \quad (6)$$

Since equation (6) does not have an unique solution, it is reasonable to introduce the condition

$$\alpha\beta / [(\alpha + \beta)^2(\alpha + \beta + 1)] = V_0, \quad (7)$$

since Var(X) = Var(1 - X). In the above, V_0 is a pre-specified quantity. From (7), we may readily observe that a solution to α exists provided that

$$V_0 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \le \frac{1}{4(\alpha+\beta+1)} < \frac{1}{4}\ .$$

However, equation (6) is not easy to solve analytically, even if it is expressed as a function of one unknown quantity, say α , in view of the above consideration. In Table 1 we display solutions to α and β satisfying (6) for selected values of V_0 and θ_0. Define

$$\Delta(\theta; \alpha, \beta) = E[|\ X - \theta|], X \sim B(x; \alpha, \beta). \quad (8)$$

Since

$$\Delta(\theta; \alpha, \beta) = \Delta(1 - \theta; \beta, \alpha), \quad (9)$$

in Table 1, we present values θ up to 0.5. Moreover, we also display the efficiency ratio $Q \times 100$, where $Q = 1/E(AI)$. The smaller the value of Q the better the degree of apportionment will be attained. In Figure 1, we display $Q \times 100$ values vs. V_0 for selected values of θ_0.

The purpose of Table 1 is to demonstrate the specification of the parameters of the beta distribution for X when the target (θ) is specified, and we maximize the AI in an averaged sense for a given value of Var(X). Some of the findings are displayed in Figure 1.

Figure 1 shows that for a specified target value θ_0 of θ, Q (reciprocal of averaged AI) increases in V_0. In other words, if we seek to be liberal (by allowing a larger variation in the X distribution by taking a higher value of Var(X)), then we will tend to achieve a poorer apportionment on an average. From the figure we can determine the extent of variation to be allowed in the X distribution to meet any specific value of the averaged AI.

Table 1. Values of α, β and $Q \times 100$ subject to (6) for a given θ and Var(X).

Var(X) = 0.01										
θ	.05	.1	.15	.2	.25	.3	.35	.4	.45	.5
α	.60	1.24	2.16	3.34	4.71	6.21	7.76	9.29	10.73	12.0
β	6.31	8.69	10.77	12.39	13.48	14.05	14.13	13.77	13.04	12.0
100Q	107.4	108.2	108.5	108.6	108.7	108.7	108.8	108.8	108.8	108.8
Var(X) = 0.05										
θ	.05	.1	.15	.2	.25	.3	.35	.4	.45	.5
α	.27	.41	.56	.74	.94	1.16	1.38	1.60	1.81	2.0
β	1.42	1.71	1.93	2.11	2.23	2.30	2.30	2.25	2.15	2.0
100Q	117.6	119.7	121.0	121.8	122.3	122.7	122.9	123.0	123.1	123.1
Var(X) = 0.10										
θ	.05	.1	.15	.2	.25	.3	.35	.4	.45	.5
α	.18	.25	.31	.37	.43	.50	.57	.63	.70	.75
β	.60	.69	.75	.79	.82	.84	.84	.82	.79	.75
100Q	129.0	132.0	134.0	135.4	136.5	137.3	137.9	138.3	138.5	138.5
Var(X) = 0.20										
θ	.05	.1	.15	.2	.25	.3	.35	.4	.45	.5
α	.073	.084	.092	.099	.10	.11	.11	.12	.12	.13
β	.11	.12	.13	.13	.13	.13	.13	.13	.13	.13
100Q	162.5	166.3	168.8	170.8	172.3	173.4	174.3	174.9	175.2	175.3

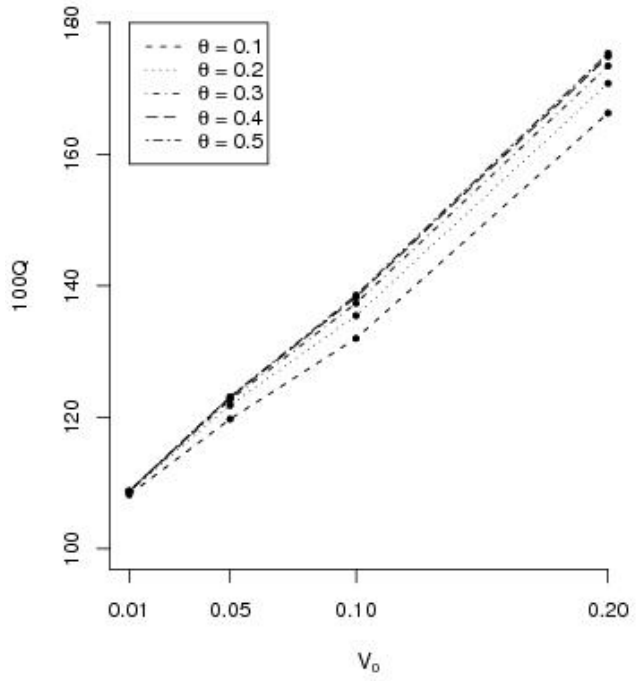


Figure 1. Graph showing $Q \times 100$ vs. V_0 for some selected values of θ in the case of two locations.

2.2. The case of three locations

We now pass on to a discussion for three locations with relative demands to be denoted by θ_1, θ_2 and θ_3 subject to the total being 1. We will also denote the corresponding random relative outputs by the variables X_1, X_2 and X_3, respectively. Note that X_1, X_2 and X_3 are non-negative random variables subject to the normalizing constraint $X_1 + X_2 + X_3 = 1$. Therefore, as a natural generalization of the beta distribution for two locations, we here adopt 3-variate singular Dirichlet distribution to describe the joint distribution of the Xs. We take the parameters of the Dirichlet distribution as α_1, α_2 and α_3 and denote the distribution by Dir[α_1, α_2, α_3] and assume (X_1, X_2) to follow this distribution, unless otherwise stated. For the sake of completeness, we display the density underlying Dir[α_1, α_2, α_3] below.

$$f(x_1, x_2; \alpha_1, \alpha_2, \alpha_3) = x_1^{\alpha_1 - 1} x_2^{\alpha_2 - 1} (1 - x_1 - x_2)^{\alpha_3 - 1} / D(\alpha_1, \alpha_2, \alpha_3),$$

$$0 < x_1, x_2 < x_1 + x_2 \leq 1, \tag{10}$$

where $D(\alpha_1, \alpha_2, \alpha_3) = \Gamma(\alpha_1, \alpha_2, \alpha_3) / \Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)$. Here $\Gamma(\alpha)$ refers to the standard Gamma integral defined for $\alpha > 0$ as $\int_0^\infty e^{-x} x^{\alpha - 1} dx$. See Johnson & Kotz (1972), pp 231-235; Kotz & Johnson (1982), Vol.2, pp. 386-387.

This time our requirements are fairly stringent so far as attainment of the absolute maximum of the expected AI is concerned. In the case of three locations the AI is given by the formula

$$AI = 1 - \frac{1}{2} [\, | X_1 - \theta_1 | + | X_2 - \theta_2 | + | X_3 - \theta_3 | \,] \tag{11}$$

and the expected AI is

$$E(AI) = 1 - \frac{1}{2} [E[|X_1 - \theta_1|] + E[|X_2 - \theta_2|] + E[|X_3 - \theta_3|]]. \tag{12}$$

Note that now $X_3 = 1 - X_1 - X_2$ and $\theta_3 = 1 - \theta_1 - \theta_2$ but for simplicity we remain in the notations X_3 and θ_3. To maximize (12) we need to minimize

$$\psi(\alpha) = E[|(X_1 - \theta_1)|] + E[|(X_2 - \theta_2)|] + E[|(X_3 - \theta_3)|] \quad (13)$$

We know that for any individual term above, minimization is achieved by taking the demand parameter θ_i as the median of the corresponding output distribution of X_i. However, it is *not* possible to attain this feature for all three terms simultaneously. To see this, we refer to Statement II in the Appendix.

Note that in a Dirichlet distribution, each marginal distribution is beta. Hence in order for all the terms to be simultaneously minimized to attain the least possible value corresponding to the median in each case, we must have

(i) $0.5 = \int_0^{\theta_1} f(x_1; \alpha_1, \alpha_2 + \alpha_3)\, dx_1 = \Delta(\theta_1; \alpha_1, t_1\alpha_1)$, where $t_1 = \dfrac{\alpha_2 + \alpha_3}{\alpha_1}$;

(ii) $0.5 = \int_0^{\theta_2} f(x_2; \alpha_2, \alpha_1 + \alpha_3)\, dx_2 = \Delta(\theta_2; \alpha_2, t_2\alpha_2)$, where $t_2 = \dfrac{\alpha_1 + \alpha_3}{\alpha_2}$;

(iii) $0.5 = \int_0^{\theta_3} f(x_3; \alpha_3, \alpha_1 + \alpha_2)\, dx_3 = \Delta(\theta_3; \alpha_3, t_3\alpha_3)$, where $t_3 = \dfrac{\alpha_1 + \alpha_2}{\alpha_3}$.

Now, appealing to Statement II in the Appendix, we must have

$$\theta_1(\alpha_2 + \alpha_3) < (1 - \theta_1)\alpha_1;\ \theta_2(\alpha_1 + \alpha_3) < (1 - \theta_2)\alpha_2;\ \theta_3(\alpha_1 + \alpha_2) < (1 - \theta_3)\alpha_3. \quad (14)$$

This pre-supposes that each θ_i is less than 0.5, which will be assumed throughout.

Summing over all the conditions and re-writing the inequality, we obtain

$$(\theta_1 + \theta_2 + \theta_3)(\alpha_1 + \alpha_2 + \alpha_3) < (\alpha_1 + \alpha_2 + \alpha_3),$$

i.e. $(\theta_1 + \theta_2 + \theta_3) < 1$ which is a contradiction. Therefore, we must have equality in each of the requirements above. This means that

$$0.5 = \int_0^{\theta} f(x; \alpha, t\alpha)\, dx \quad (15)$$

is to be satisfied for a finite α while $t = (1 - \theta)/\theta$. This is again a contradiction, as indicated in the Appendix.

Simultaneous minimization of all three terms in (13) to respective absolute minimum must therefore be ruled out. From now onwards, we assume $\theta_1 \le \theta_2 \le \theta_3$ without any loss of generity. However, a unique choice of α_is may be made by selecting them in the ratio of the θ_is and by equating the *highest marginal variance* of the X_is to a given quantity V_0. In other words, we may start with $[\alpha, a\alpha, b\alpha]$, where $a = \theta_2 / \theta_1$, $b = \theta_3 / \theta_1$ $(1 \le a \le b)$, and seek to choose α using the variance requirement V_0 on the largest of $V(X_1)$, $V(X_2)$ and $V(X_3)$, i.e., on $V(X_3)$. We can compute the value of AI and hence that of Q and examine its behaviour for variations in α for given (a, b), i.e. for given θ_1, θ_2 and θ_3. In Tables 2, 3 and 4 we show the values of Q x 100 and $\psi(\alpha)$ in (13) for different values of θ_1, θ_2, θ_3 as a function of V_0 (or α).

Tables 2, 3 and 4 and also Figure 2 show that for higher values of V_0, equal-demand distribution seems to yield less satisfactory results. We should therefore look at this case more carefully. In Table 5 we compare the relative ratios of Q_i and Q_j by using the measure RR(i, j) = $\dfrac{|Q_i - Q_j|}{\frac{1}{2}(Q_i + Q_j)} \times 100\%$, where indices i and j

correspond to the settings in Tables 2, 3 and 4, respectively. Table 5 and also Figure 2 reveal that there is little difference among the realized Q values for different demand distributions whenever V_0 is appreciably small. The comparison also shows that settings 1 and 2 likewise differ little, but a slight difference is obtained for large values of V_0. For settings 1 and 3, and also for settings 2 and 3, the difference is fairly large except for notably small values of V_o.

The values of $\Psi(\alpha)$ and $Q \times 100$ for different values of θ_1, θ_2, θ_3 as a function of V_0 (or α).

Table 2. $\theta_1 = 0.17$, $\theta_2 = 0.38$, $\theta_3 = 0.45$

$$a = \frac{\theta_2}{\theta_1} = 2.23529,\ b = \frac{\theta_3}{\theta_1} = 2.64706$$

V_0	α	$\psi(\alpha)$	100Q
.01	4.04	0.219	112.3
.02	1.93	0.312	118.5
.03	1.23	0.385	123.8
.04	0.88	0.447	128.8
.05	0.67	0.504	133.6
.075	0.39	0.625	145.5
.10	0.25	0.733	157.8
.125	0.17	0.831	171.1
.15	0.11	0.923	185.7
.175	0.070	1.011	202.1
.20	0.040	1.095	221.0

Table 3. $\theta_1 = 0.1$, $\theta_2 = 0.45$, $\theta_3 = 0.45$

$$a = \frac{\theta_2}{\theta_1} = 4.5,\ b = \frac{\theta_3}{\theta_1} = 4.5$$

V_0	α	$\psi(\alpha)$	100Q
.01	2.38	0.209	111.6
.02	1.14	0.296	117.4
.03	0.73	0.364	122.3
.04	0.52	0.422	126.8
.05	0.47	0.474	131.1
.075	0.23	0.588	141.6
.10	0.15	0.687	152.4
.125	0.098	0.778	163.7
.15	0.065	0.864	176.1
.175	0.041	0.946	189.8
.20	0.024	1.025	205.2

Table 4. $\theta_1 = \frac{1}{3},\ \theta_2 = \frac{1}{3},\ \theta_3 = \frac{1}{3}$

$a = \frac{\theta_2}{\theta_1} = 1.0,\ b = \frac{\theta_3}{\theta_1} = 1.0.$

V_0	α	$\psi(\alpha)$	100Q
.01	7.07	0.215	113.7
.02	3.37	0.345	120.8
.03	2.14	0.426	127.1
.04	1.52	0.496	133.0
.05	1.15	0.560	138.8
.075	0.65	0.699	153.9
.10	0.41	0.823	170.0
.125	0.26	0.938	188.3
.15	0.16	1.045	209.5
.175	0.090	1.148	234.8
.20	0.037	1.247	265.8

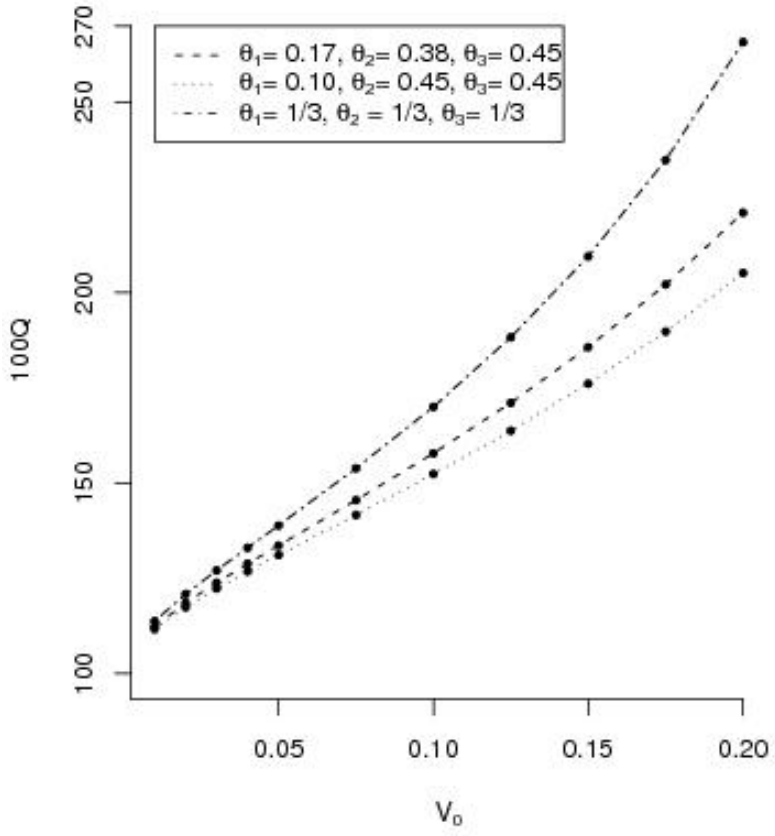


Figure 2. Graph displaying $Q \times 100$ vs. V_0 for different settings in Example 2.1.

Table 5. Relative ratio of Q values defined as $RR(i,j) = \frac{|Q_i - Q_j|}{\frac{1}{2}(Q_i - Q_j)} \times 100\%$, where i and j refer to two different demand distributions.

V_0	RR(1,1)	RR(1,3))	RR(2,3)
.01	0.60	1.26	1.86
.02	0.95	1.95	2.89
.03	1.27	2.58	3.85
.04	1.58	3.21	4.79
.05	1.89	3.85	5.74
.075	2.69	5.60	8.29
.10	3.51	7.43	10.94
.125	4.38	9.58	13.94
.15	5.30	12.06	17.33
.175	6.30	14.96	21.21
.20	7.42	18.41	25.74

The computations in this section were carried out using Mathematica (Wolfram, 1999) and R software (Venables & Ripley 2002) environments. The software code is available from the authors on request.

3. DISCUSSION

In this paper we have initiated a statistical study of the Apportionment Index (AI) by considering it as a random variable and by seeking maximization of its average value. We have developed the necessary theory and computational aspects for this problem in the case of two and three locations. Our results rely on a specification of the distribution of the underlying random variable(s) which is taken to be beta (Dirichlet). The general case is yet to be taken up. There is also scope for a Bayesian analysis of this problem by considering the target parameters (θ) to be random and taking appropriate priors for them.

4. ACKNOWLEDGMENTS

This work was financed by the Academy of Finland under the project entitled "Advanced Methods for Computer-Aided Bucking of Scots Pine" (project 73691). We are grateful to an anonymous referee for critically reading the manuscript and offering helpful suggestions which improved the presentation of our ideas.

REFERENCES

BERGSTRAND, K.G. (1990): "(Bucking to demand, the way to fulfil sawmill objectives) Fördelningsaptering-ett sätt att tillgodose sägverksönskemål". Skogforsk Resultat Nr 11/1990.

JOHNSON, N.L. and S. KOTZ (1972): **Distributions in Statistics: Continuous Multivariate Distributions**. Wiley & Sons. New York.

JOHNSON, N.L.; S. KOTZ and N. BALAKRISHNAN (1995): **Continuous Univariate Distributions**, 2, 2nd ed. Wiley & Sons. New York

KIRKKALA, A., **et al**. (2000): "(Penalty segmented demand distribution for evaluating bucking outcome and calculating distribution level) Sakkosegmentoitu tavoitejakauma apteeraustuloken arvioinnissa ja jakauma-asteen laskennassa", Metsätieteen aikakauskirja 1/2000: 59:61.

KIVINEN, V.P.; J. UUSITALO and T. NUMMI (2003): "Measuring the Goodness of the Bucking Outcome". Proceedings of Second Forest Engineering Conference (Arbetsrapport från Skogforsk 537, 87-95) Växjö, Sweden.

KOTZ, S. and N.L. JOHNSON Edit. (1982): **Encyclopedia of Statistical Sciences**,1-2. Wiley Interscience. New York.

MALINEN, J. and T. PALANDER (2004): "Metrics for distribution similarity applied to the bucking to demand procedure", **International Journal of Forest Engineering**, 15: 33-40.

NUMMI, T.; B. SINHA and L. KOSKELA (2004): "Statistical Properties of the Apportionment Degree and Alternative Measures in Bucking Outcome", To appear in **Investigación Operacional** 26.

VENABLES, W.N. and B.D. RIPLEY (2002): **Modern Applied Statistics with S**, 4th ed.: Springer,. New York.

WOLFRAM, S. (1999): **The Mathematica Book**, 4th ed. Wolfram Media/Cambridge University Press, Cambridge.

A P P E N D I X

STATEMENT I

Define

$$\Delta(r;\alpha,t\alpha)=\frac{\int_0^r x^{\alpha-1}(1-x)^{t\alpha-1}dx}{\int_0^1 x^{\alpha-1}(1-x)^{t\alpha-1}dx}\,. \tag{16}$$

Then $\Delta(r;\alpha,t\alpha)\uparrow$ in t, $\forall\alpha>0$, $\forall r\in(0,1)$.

Proof.

Taking the first derivative and requiring it to be positive, we end up, after simplification, with the inequality

$$\left(\int_0^r x^{\alpha-1}(1-x)^{t\alpha-1}\alpha\left(-\log(1-x)\right)dx\right)\left(\int_0^1 x^{\alpha-1}(1-x)^{t\alpha-1}dx\right)<$$

$$<\left(\int_0^r x^{\alpha-1}(1-x)^{t\alpha-1}dx\right)\left(\int_0^1 x^{\alpha-1}(1-x)^{t\alpha-1}\alpha\left(-\log(1-x)\right)dx\right), \tag{17}$$

which is equivalent to

$$\sum_{j=1}^{\infty}\frac{1}{j}\left(\int_0^r x^{\alpha+j-1}(1-x)^{t\alpha-1}dx\right)\left(\int_0^1 x^{\alpha-1}(1-x)^{t\alpha-1}dx\right)<\sum_{j=0}^{\infty}\frac{1}{j}\left(\int_0^r x^{\alpha-1}(1-x)^{t\alpha-1}dx\right)\left(\int_0^1 x^{\alpha+j-1}(1-x)^{t\alpha-1}dx\right). \tag{18}$$

Then, comparing term by term, we require

$$\left(\int_0^r x^{\alpha+j-1}(1-x)^{t\alpha-1}dx\right)\left(\int_0^1 x^{\alpha-1}(1-x)^{t\alpha-1}dx\right)<\left(\int_0^r x^{\alpha-1}(1-x)^{t\alpha-1}dx\right)\left(\int_0^1 x^{\alpha+j-1}(1-x)^{t\alpha-1}dx\right), \tag{19}$$

which can be expressed as

$$\left(\int_0^r x^{\alpha+j-1}(1-x)^{t\alpha-1}dx\right)\left(\int_0^r x^{\alpha-1}(1-x)^{t\alpha-1}dx+\int_r^1 x^{\alpha-1}(1-x)^{t\alpha-1}dx\right)<$$

$$< \left(\int_0^r x^{\alpha-1}(1-x)^{t\alpha-1}dx \right) \left(\int_0^r x^{\alpha+j-1}(1-x)^{t\alpha-1}dx + \int_r^1 x^{\alpha+j-1}(1-x)^{t\alpha-1}dx \right). \quad (20)$$

We can now simplify the above and demand:

$$\left(\int_0^r x^{\alpha+j-1}(1-x)^{t\alpha-1}dx \right) \left(\int_r^1 x^{\alpha-1}(1-x)^{t\alpha-1}dx \right) < \left(\int_0^r x^{\alpha-1}(1-x)^{t\alpha-1}dx \right) \left(\int_r^1 x^{\alpha+j-1}(1-x)^{t\alpha-1}dx \right). \quad (21)$$

Note next that on the left-hand side, $x^j \le r^j$, while on the right, $x^j \le r^j$, and this is true for all j = 0, 1, 2,... Therefore

$$\text{left-side integral} < r^j \left(\int_0^r x^{\alpha-1}(1-x)^{t\alpha-1}dx \right) \left(\int_0^1 x^{\alpha-1}(1-x)^{t\alpha-1}dx \right),$$

while

$$\text{right-side integral} > r^j \left(\int_0^r x^{\alpha-1}(1-x)^{t\alpha-1}dx \right) \left(\int_r^1 x^{\alpha-1}(1-x)^{t\alpha-1}dx \right).$$

Hence the claim is settled.

It thus follows that $\Delta(r; \alpha, t_1\alpha) \le \Delta(r; \alpha, \alpha) \le \Delta(r; \alpha, t_2\alpha)$, $\forall t_1 < 1 < t_2$, $\forall \alpha > 0$, $\forall r \in (0,1)$.

STATEMENT II

With the Δ function defined as in Statement I, $\Delta(r; \alpha, t\alpha) = 0.5$ is possible only when $t \lessgtr (1 - r)/r$, according as $r \lessgtr 0.5$ whatever be the value of α.

A satisfactory analytical proof of Statement II has so far eluded us. However, we have carried out extensive numerical computations and our results support the claim. In Table 6 we display some of the computations (see also Figure 3).

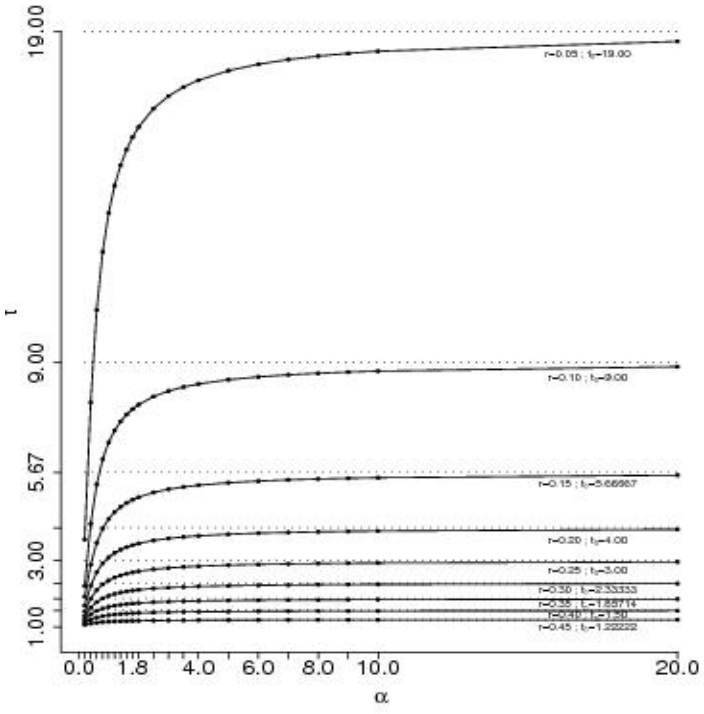


Figure 3.
Graph showing $t(\alpha)$ as a function of α satisfying the equation

$$\int_0^r \frac{1}{B(\alpha, t\alpha)} x^{\alpha-1}(1-x)^{t\alpha-1}dx = 0.5$$

for some selected values of r

(r = 0.05, 0.1,…,0.45).

The values of $t_0(r) = \frac{1}{r} - 1$ are also indicated along the t-axis.

Table 6. Table showing selected values of r and t for different values of α. The values of $t_0(r) = \frac{1}{r} - 1$ are also shown in the table.

$$\int_0^r \frac{1}{B(\alpha, t\alpha)} x^{\alpha-1}(1-x)^{t\alpha-1}dx = 0.5$$

α	r									
		.05	.10	.15	.20	.25	.30	.35	.40	.45
0.2		3.642	2.240	1.919	1.647	1.465	1.331	1.226	1.139	1.065
0.4		7.787	4.126	2.883	2.247	1.855	1.585	1.386	1.231	1.105
0.6		10.583	5.310	3.544	2.655	2.117	1.754	1.491	1.290	1.131
0.8		12.340	6.068	3.973	2.923	2.290	1.866	1.560	1.329	1.147
1.0		13.513	6.579	4.265	3.106	2.409	1.943	1.609	1.357	1.159
1.2		14.344	6.942	4.474	3.238	2.496	2.000	1.645	1.377	1.168
1.4		14.960	7.213	4.629	3.337	2.560	2.042	1.671	1.392	1.175
1.6		15.433	7.421	4.750	3.413	2.611	2.075	1.692	1.405	1.180
1.8		15.809	7.587	4.845	3.474	2.651	2.102	1.709	1.414	1.184
2.0		16.113	7.721	4.923	3.524	2.684	2.123	1.723	1.422	1.188
2.5	t	16.670	7.967	5.066	3.615	2.744	2.163	1.748	1.437	1.194
3.0		17.048	8.134	5.163	3.677	2.785	2.190	1.766	1.447	1.199
3.5		17.320	8.255	5.233	3.722	2.815	2.210	1.778	1.454	1.202
4.0		17.526	8.346	5.286	3.755	2.837	2.225	1.788	1.460	1.204
5.0		17.817	8.475	5.361	3.803	2.869	2.246	1.801	1.465	1.208
6.0		18.011	8.561	5.411	3.836	2.891	2.260	1.810	1.473	1.210
7.0		18.151	8.623	5.447	3.859	2.906	2.271	1.817	1.477	1.212
8.0		18.256	8.670	5.474	3.876	2.918	2.278	1.822	1.479	1.213
9.0		18.338	8.706	5.495	3.890	2.927	2.284	1.826	1.482	1.214
10.0		18.404	8.735	5.512	3.901	2.934	2.289	1.829	1.483	1.215
20.0		18.701	8.867	5.589	3.950	2.967	2.311	1.843	1.492	1.219
	t_0	19.00	9.00	5.66667	4.00	3.00	2.33333	1.85714	1.50	1.22222