NOTE / NOTE

# A note on the relationship between the quadratic mean stand diameter and harmonic mean basal area under size-biased distribution theory

**Jeffrey H. Gove**

**Abstract:** This note seeks to extend the utility of size-biased distribution theory as applied to forestry through two relationships regarding the quadratic mean stand diameter. First, the quadratic mean stand diameter's relationship to the harmonic mean basal area for horizontal point sampling, which has been known algebraically from early on, is proved under size-biased distribution theory. Second, a new result, which may prove most valuable in viewing the graphical representation of assumed distributions, is also derived. The results are also shown to apply to the basal area – size distribution, providing a unique duality between the two means.

**Résumé :** Cette note vise à élargir la portée de la théorie de la distribution diamétrale biaisée telle qu'elle est appliquée en foresterie par le biais de deux propriétés du diamètre moyen quadratique du peuplement. D'abord, la théorie est utilisée pour démontrer la relation qui existe entre le diamètre moyen quadratique du peuplement et la surface terrière moyenne harmonique dans le cas de l'échantillonnage horizontal par point ; une relation dont la démonstration algébrique est connue depuis longtemps. Deuxièmement, un nouveau résultat qui peut s'avérer très utile pour représenter graphiquement des distributions présumées est également dérivé de la théorie. Ces résultats sont aussi applicables à la distribution de la surface terrière par rapport au diamètre établissant ainsi une dualité unique entre le diamètre moyen et la surface terrière moyenne.

[Traduit par la Rédaction]

## Introduction

Foresters rely heavily on distributions, either empirical or in assumed parametric form, in much of their data analysis work when quantities based on tree diameter are reported. A number of mean values, some specific to forestry, can be calculated from these distributions. The interrelationships of three means in particular under size-biased distribution theory were discussed by Gove and Patil (1998). Of the pool available, the quadratic mean stand diameter seems to be the most commonly used. This is perhaps because of its dual identity as the diameter of the tree of mean basal area; since basal area is among the most useful variables in forest management, silviculture, and modeling, its utility is not surprising. Two interesting properties associated with the quadratic mean stand diameter are described herein with respect to assumed size-biased distributions.

Size-biased probability distributions arise as a special case of weighted distribution theory (Patil and Rao 1977) in at least two common contexts in forest mensurational work. First, the relationship between diameter and basal area yields the basal area – size distribution as a reapportionment of the usual diameter–frequency probability density (Gove and Patil 1998). Second, when sampling with probability proportional to size under horizontal point sampling (HPS), the distribution of the sampled tree diameters is related to the underlying per-unit area diameter distribution through this theory (Gove 2000; Van Deusen 1986). In both instances, the size-biased density is of order $\alpha = 2$. Briefly, if a random variable such as tree diameter ($D$) at breast height (DBH) has probability density $f(d;\boldsymbol{\theta})$, then the size-biased probability density is given as

$$f_\alpha^*(d;\boldsymbol{\theta}) = \frac{d^\alpha f(d;\boldsymbol{\theta})}{\mu'_\alpha}$$

where $\mu'_\alpha = \int d^\alpha f(d;\boldsymbol{\theta})\,\delta d$,[1] and $\boldsymbol{\theta}$ is the parameter vector. A graphical example of this relationship is shown in Fig. 1, where

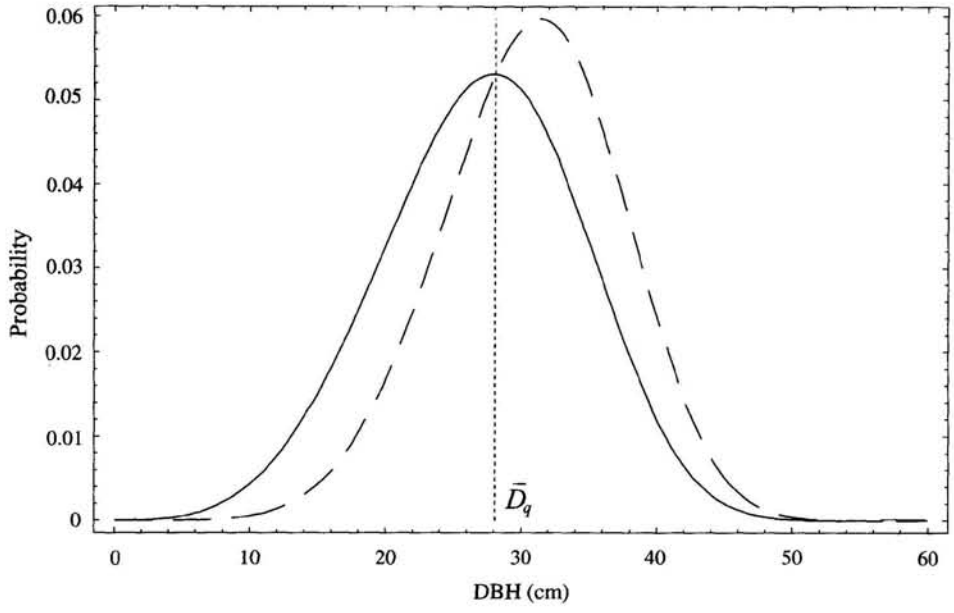
Received 30 April 2002. Accepted 29 January 2003. Published on the NRC Research Press Web site at http://cjfr.nrc.ca/ on 17 July 2003.

**J.H. Gove.** USDA Forest Service, Northeastern Research Station, 271 Mast Road, P.O. Box 640, Durham, NH 03824, U.S.A. (e-mail: jgove@fs.fed.us).

[1]The differential operator $\delta$ is used in diameter-related integrals to avoid confusion with $d$, the realization of the random variable $D$ for diameter.

 doi: 10.1139/X03-043 © 2003 NRC Canada

**Fig. 1.** DBH-frequency distribution $f(d;\boldsymbol{\theta})$ (solid line) and size-biased form $f_2^*(d;\boldsymbol{\theta})$ (broken line) for a stand with basal area and number of trees per hectare $B = 45.91$ m$^2$·ha$^{-1}$ and $N = 741.3$. Here, $f(d;\boldsymbol{\theta})$ is assumed to be a two-parameter Weibull ($\boldsymbol{\theta} = (\gamma, \beta)$) with shape and scale parameters $\gamma = 4.1709$ and $\beta = 29.8384$, respectively. The quadratic mean stand diameter is $\bar{D}_q = 28.08$ cm.




$\alpha = 2$. In this example a two-parameter Weibull has been used for illustration with

$$f(d;\boldsymbol{\theta}) = \frac{\gamma}{\beta}\left(\frac{d}{\beta}\right)^{\gamma-1} \mathrm{e}^{-(d/\beta)^{\gamma}}$$

and (Gove and Patil 1998)

$$f_2^*(d;\boldsymbol{\theta}) = \frac{d^2\frac{\gamma}{\beta}\left(\frac{d}{\beta}\right)^{\gamma-1} \mathrm{e}^{-(d/\beta)^{\gamma}}}{\beta^2\Gamma(2/\gamma+1)}$$

where for both densities $d$, $\gamma$ and $\beta > 0$, with scale and shape parameters $\boldsymbol{\theta} = (\gamma, \beta) = (4.1709, 29.8384)$, respectively, and $\Gamma(\cdot)$ is the gamma function. Because of the intrinsic relationship between the basal area – size distribution and the distribution of diameters from a HPS tally, $f_2^*(d;\boldsymbol{\theta})$ could, in general, represent either density in this figure. Note that the Weibull is used simply for illustration; in what follows no parametric family of distributions is assumed.

The compatibility between the size-biased distributional form ($f_2^*(d;\boldsymbol{\theta})$) and the original diameter distribution ($f(d;\boldsymbol{\theta})$) in either setting relies on the fact that $\boldsymbol{\theta}$ is a common shared parameter set. Given that $\boldsymbol{\theta}$ is common to both distributional forms, compatibility can be shown in a variety of ways (Gove 2000; Gove and Patil 1998; Van Deusen 1986). In this note, the argument for compatibility of the size-biased forms is further supported via a relationship involving the quadratic mean stand diameter and the harmonic mean. In addition, a new basic result linking the quadratic mean stand diameter to the size-biased form is proven. Both results are shown to be independent of the distributional family assumed in parameterization.

## The harmonic mean – quadratic mean stand diameter relationship

One place in forestry where the harmonic mean has been shown to be applicable is in horizontal point sampling (HPS). Buckingham (1969) and Bitterlich (1984, p. 40) review this application of the harmonic mean and describe a result that is attributable to Hirata (1956) and Bitterlich (1957), namely that the harmonic mean basal area ($\bar{B}_\mathrm{h}$) for the trees tallied on $m$ HPS points is equivalent to the arithmetic mean basal area ($\bar{B}_\mathrm{a}$) for the trees in the sampled population. If $B$ and $N$ are the usual HPS estimates of the basal area and number of trees per hectare, respectively, for a given area taken from $m$ points, then the arithmetic mean basal area is

$$\begin{aligned}\bar{B}_\mathrm{a} &= \frac{B}{N}\\ &= \kappa\bar{D}_\mathrm{q}^2\end{aligned}$$

where $\kappa$ is the conversion from square centimetres to square metres, and $\bar{D}_\mathrm{q}$ is the quadratic mean stand diameter — both quantities are considered estimates due to sampling. Now, if $N^*$ trees were actually tallied on these same $m$ HPS points, then the harmonic mean basal area for the tree tally is given as

$$[1] \qquad \bar{B}_\mathrm{h} = \frac{N^*}{\sum_{i=1}^{N^*}\frac{1}{b_i}}$$

where $b_i$ is the basal area for the $i$th tree sampled. The result that $\bar{B}_\mathrm{a} \equiv \bar{B}_\mathrm{h}$ follows by simple algebra; worked examples are given in Buckingham (1969), Husch et al. (1982, p. 263), and Bitterlich (1984, p. 191).

Inasmuch as size-biased distribution theory is applicable to HPS, this result should also be able to be proved using the

©2003 NRC Canada

respective probability densities. Indeed, if the distribution of HPS diameters is given as $f_2^*(d;\boldsymbol{\theta})$, then the equivalence of the preceding formulas can be shown when written in terms of the diameter distribution for the HPS, or Bitterlich (as Hirata (1956) referred to it), population. In general, for a size-biased distribution of order $\alpha = 1$, the harmonic mean of the distribution is given as $E[D^{*-1}]^{-1}$ (Mahfoud and Patil 1982). By analogy, the harmonic mean basal area is therefore denoted as $E[1/\kappa D^{*2}]^{-1}$. Hence, first note that

$$[2] \qquad E\left[\frac{1}{\kappa D^{*2}}\right] = \int \frac{1}{\kappa d^2} f_2^*(d;\boldsymbol{\theta})\,\delta d = \int \frac{1}{d^2} \times \frac{d^2 f(d;\boldsymbol{\theta})}{\kappa \mu_2'}\,\delta d = \frac{1}{\kappa \mu_2'} \int f(d;\boldsymbol{\theta})\,\delta d = \frac{1}{\kappa \mu_2'}$$

Then, the harmonic mean basal area is given as

$$\frac{1}{E\left[\frac{1}{\kappa D^{*2}}\right]} = \kappa \mu_2' = \kappa \int d^2 f(d;\boldsymbol{\theta})\,\delta d = \kappa \bar{D}_{\mathrm{q}}^2$$

The last step follows, since the second raw moment of the population per-unit area diameter distribution, $f(d;\boldsymbol{\theta})$, is the square of the quadratic mean stand diameter $\bar{D}_{\mathrm{q}}$, regardless of the parametric form (Ek et al 1975; Burk and Newberry 1984; Gove and Patil 1998), and the result has been shown. Thus, weighted distribution theory preserves the intrinsic relationship between HPS tally and population data and yields an alternative proof in terms of probability densities for this well-know result in forestry. This result is also consistent with existing statistical theory, since, in general, it is well known that the harmonic mean of the size-biased distribution of order $\alpha = 1$ is the mean of the original unweighted distribution (Mahfoud and Patil 1982; Jones 1990). This result, however, extends these findings to a simple function of the size-biased distribution of order $\alpha = 2$ that is undoubtedly specific to forestry.

Because of the relationship between the HPS diameter distribution and the basal area – size distribution, the proof above also applies to the latter. It is straightforward to verify this analytically by noting that the harmonic mean basal area can be recast in terms of the basal area – size distribution, namely

$$[3] \qquad \bar{B}_{\mathrm{h}} = \frac{B}{\sum_{i=1}^{n} \frac{B_i}{b_i}}$$

where $B_i$ is the basal area per-unit area in the $i$th diameter class, and $b_i = \kappa d_i$ is the class midpoint basal area for $i = 1, \ldots, n$. It is important to notice here that the probability density in the basal area – size distribution is in terms of basal area, not frequency; therefore, the tree frequencies in [1] must be replaced with the respective basal area weights $B$ and $B_i$. It is straightforward to show that the denominator in [3] reduces to $N$. Thus, the harmonic mean of the basal area – size distribution is also equal to $\bar{B}_{\mathrm{a}}$.

## A second relationship

The result of the preceding section should not be overly surprising. Size-biased distribution theory has been shown to provide a general theoretical framework for linking the underlying theoretical population diameter distribution with both the basal area – size distribution and the distribution of diameters from the Bitterlich population. Thus, relationships that are found to hold in the general empirical distribution should also hold under this more general theoretical framework. However, one further interesting result that evidently can not be arrived at without resorting to size-biased distribution theory is the following.

Consider again the graph in Fig. 1 showing both $f(d;\boldsymbol{\theta})$ and $f_2^*(d;\boldsymbol{\theta})$. The point of intersection for the two densities is seen to be at $\bar{D}_{\mathrm{q}}$ given the data that generated it. A relevant question then becomes, is this always the case, regardless of the shape of the underlying distribution? To answer this question, set the two densities equal and solve for the point of intersection in terms of the original random variable, namely

$$f(d;\boldsymbol{\theta}) = f_2^*(d;\boldsymbol{\theta}) = \frac{d^2 f(d;\boldsymbol{\theta})}{\mu_2'}$$

or

$$\frac{f(d;\boldsymbol{\theta})}{f(d;\boldsymbol{\theta})} = \frac{d^2}{\mu_2'}$$

which implies that $d^2 = \mu_2'$. But we have already seen that the second raw moment of the original DBH-fequency distribution, $\mu_2'$, is the quadratic mean stand diameter squared. Thus, taking the square root of both sides, we have finally that the point of intersection is at $d = \bar{D}_{\mathrm{q}}$. Because no distributional family has been assumed in this result (e.g., Weibull, beta, etc.), it holds regardless of the assumed form for $f(d;\boldsymbol{\theta})$ and thus $f_2^*(d;\boldsymbol{\theta})$.

The results have been presented here in terms of the respective probability densities for generality sake. However, it is again straightforward to show that this result holds for both the HPS and basal area – size distributions when appropriate scale factors have been applied to expand the densities to the respective per-unit area quantities. Thus, the intersection of the size-biased distribution of order $\alpha = 2$ and the original DBH-frequency distribution always occurs at the quadratic mean stand diameter.

As an aside, a related result can similarly be proven for the size-biased distribution of order $\alpha = 1$, namely $f_1^*(d;\boldsymbol{\theta})$ will always cross $f(d;\boldsymbol{\theta})$ at the arithmetic mean diameter $d = \mu_1'$. This result corresponds to the results on the first-order harmonic mean cited earlier (Mahfoud and Patil 1982; Jones 1990). Undoubtedly, analogous results for higher order size-biased distributions (i.e., $\alpha > 2$) also exist, although their usefulness in forestry is questionable.

©2003 NRC Canada

## Concluding remarks

The proofs in this note have shown that size-biased distribution theory provides results that are consistent with previous findings on the relationship between certain harmonic and quadratic means. In addition, a new result has been identified, constraining the crossing point of $f(d;\boldsymbol{\theta})$ and its associated size-biased counterpart.

An interesting consequence of these findings is that we can consider the theory developed in terms of a duality paradigm. Since the quadratic mean stand diameter (through $\bar{B}_a$) and the harmonic mean basal area have been shown to be equivalent under size-biased distributions of order $\alpha = 2$, we can think in terms of each of the two means "belonging" to the respective probability density. It is natural to consider the quadratic mean stand diameter as belonging to the usual diameter-frequency distribution, since this is the distribution that it is normally calculated from. However, the definitions of the harmonic mean basal area are always given in terms of the size-biased probability density. Thus we can similarly consider the harmonic mean basal area as belonging to the respective size-biased form. The duality established, therefore, is that regardless of which density one chooses to work with, either mean version will always provide the intersection of the two densities.

Plotting the two appropriately scaled densities will always give a correct graphical representation of $\bar{D}_q$ (or $\bar{B}_h$) as in Fig. 1. However, these results assume that the empirical distribution has not been truncated based on merchantability limits or the like. If a complete density is fitted to truncated data, the empirical truncated estimate of $\bar{D}_q$ will undoubtedly be different than that given by the intersection of the full densities.

## References

Bitterlich, W. 1957. Harmonische mittel in winkelzählproben. Allg. Forstztg. Vol. 68.

Bitterlich, W. 1984. The relascope idea: relative measurements in forestry. 1st ed. Commonwealth Agricultural Bureaux, Slough, U.K.

Buckingham, F.M. 1969. The harmonic mean in forest mensuration. For. Chron. **45**: 104–106.

Burk, T.E., and Newberry, J.D. 1984. A simple algorithm for moment-based recovery of Weibull distribution parameters. For. Sci. **30**: 329–332.

Ek, A.R., Issos, J.N., and Bailey, R.L. 1975. Solving for Weibull diameter distribution parameters to obtain specified mean diameters. For. Sci. **21**: 290–292.

Gove, J.H. 2000. Some observations on fitting assumed diameter distributions to horizontal point sampling data. Can. J. For. Res. **30**: 521–533.

Gove, J.H., and Patil, G.P. 1998. Modeling the basal area-size distribution of forest stands: a compatible approach. For. Sci. **44**: 285–297.

Hirata, T. 1956. Harmonic means in Bitterlich's sampling. University of Tokoyo, Tokoyo. For. Misc. Inf. 11.

Husch, B., Miller, C.I., and Beers, T.W. 1982. Forest mensuration. 3rd ed. John Wiley and Sons, New York.

Jones, M.C. 1990. The relationship between moments of truncated and original distributions plus some other simple structural properties of weighted distributions. Metrika, **37**: 233–243.

Mahfoud, M., and Patil, G.P. 1982. On weighted distributions. *In* Statistics and probability: essays in honor of C.R. Rao. *Edited by* G. Kallianpur, P.R. Krishnaiah, and J.K. Ghosh. North-Holland, New York. pp. 479–492.

Patil, G.P., and Rao, C.R. 1977. The weighted distributions: a survey of their applications. *In* Applications of statistics. *Edited by* P.R. Krishnaiah. North-Holland, New York. pp. 383–405.

Van Deusen, P.C. 1986. Fitting assumed distributions to horizontal point sample diameters. For. Sci. **32**: 146–148.

©2003 NRC Canada