

ノート (Note)

樹木の直径分布をワイブル分布へあてはめる場合におけるデータ精度の影響

伊東 宏樹*

キーワード：Stan、シミュレーション、直径分布、ベイズ推定、ワイブル分布

樹木の直径分布をモデル化するには、確率分布としてワイブル分布がしばしば用いられる (Bailey and Dell 1973, 大野 2018)。ワイブル分布は、形状パラメータと尺度パラメータの2つのパラメータで決定され、さまざまな形の直径分布を表現することが可能である (Rennolls et al. (1985) や Cheng et al. (2017) などのように、小径木のない林分では最小直径に相当するパラメータを加える場合もある)。ワイブル分布の確率密度は、 α を形状パラメータ、 σ を尺度パラメータとして次式であらわされる ($y \geq 0$, $\alpha > 0$, $\sigma > 0$)。

$$\Pr(y | \alpha, \sigma) = \frac{\alpha}{\sigma} \left(\frac{y}{\sigma}\right)^{\alpha-1} \exp\left(-\left(\frac{y}{\sigma}\right)^{\alpha}\right)$$

パラメータは一般に直径の実測値から推定されるが、この際の測定値のデータ精度は一定とは限らない。たとえば、巻尺を用いて胸高直径を測定する場合には0.1cm単位で行われることが多いが、輪尺を使用する場合には2cm括約となる場合もある。また、5cm括約などの胸高直径階分布の形でしかデータが残っていない場合もある。このようなデータを利用してワイブル分布のパラメータを推定する場合には、括約により推定値に偏りが生じないかを検証しておく必要があるだろう。今回は、比較的若齢のカンバ林を対象として、胸高直径データの測定値をワイブル分布に当てはめる場合を想定し、パラメータの推定値にデータの精度が及ぼす影響をシミュレーションにより評価した。

シミュレーションでは、若齢カンバ林に相当するような胸高直径分布を想定し、形状パラメータと尺度パラメータについて、それぞれ2と8 (平均胸高直径7.1cm、以下同じ)、4と14 (12.7cm)、6と20 (18.6cm)の3通りの設定とした (Fig. 1)。また、標本サイズは30と300の2通りとした。これらを組み合わせ、6通りの設定でシミュレーションを実行した (Table 1)。シミュレーションでは、ま

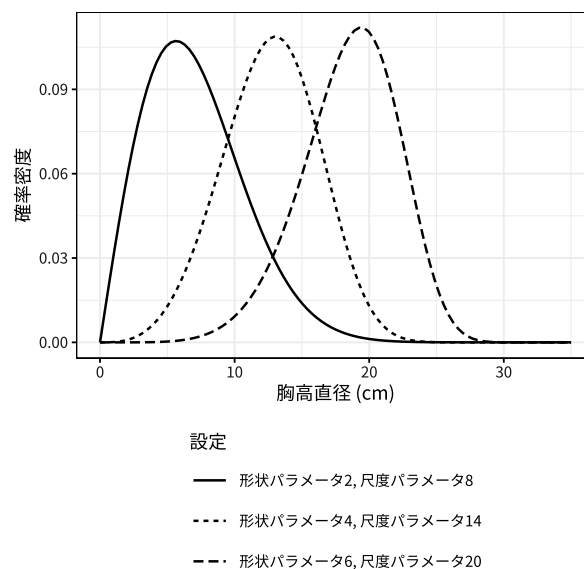


Fig. 1. シミュレーションに使用したパラメータによるワイブル分布の確率密度

Table 1. シミュレーションの設定および各設定におけるパラメータの推定結果 (事後平均値の平均±標準偏差)

シミュレーション設定			形状パラメータの推定結果				尺度パラメータの推定結果			
標本サイズ	パラメータの真値		データの精度				データの精度			
	形状	尺度	倍精度	0.1cm 単位	2cm 括約	5cm 括約	倍精度	0.1cm 単位	2cm 括約	5cm 括約
30	2	8	2.1 ± 0.3	2.1 ± 0.3	2.1 ± 0.3	2.1 ± 0.4	8.1 ± 0.8	8.1 ± 0.8	8.1 ± 0.8	8.1 ± 0.9
300	2	8	2.0 ± 0.1	2.0 ± 0.1	2.0 ± 0.1	2.0 ± 0.1	8.0 ± 0.2	8.0 ± 0.2	8.0 ± 0.2	8.0 ± 0.3
30	4	14	4.2 ± 0.6	4.2 ± 0.6	4.2 ± 0.6	4.3 ± 0.8	14.0 ± 0.7	14.0 ± 0.7	14.0 ± 0.7	14.0 ± 0.7
300	4	14	4.0 ± 0.2	4.0 ± 0.2	4.0 ± 0.2	4.0 ± 0.2	14.0 ± 0.2	14.0 ± 0.2	14.0 ± 0.2	14.0 ± 0.2
30	6	20	6.2 ± 1.0	6.2 ± 1.0	6.2 ± 1.0	6.4 ± 1.3	20.0 ± 0.6	20.0 ± 0.6	20.0 ± 0.7	20.0 ± 0.7
300	6	20	6.0 ± 0.3	6.0 ± 0.3	6.0 ± 0.3	6.0 ± 0.3	20.0 ± 0.2	20.0 ± 0.2	20.0 ± 0.2	20.0 ± 0.2

Effects of data precision in cases of fitting tree diameter distribution to Weibull distribution
Hiroki ITO*)

原稿受付：令和3年3月4日 原稿受理：令和3年3月26日

* 森林総合研究所 北海道支所 〒062-8516 北海道札幌市豊平区羊ヶ丘7

Received 4 March 2021, Accepted 26 March 2021

* Hokkaido Research Center, Forestry and Forest Products Research Institute, Toyohira-ku, Sapporo, 062-8516, JAPAN

E-mail: hiroki@ffpri.affrc.go.jp

ず、標本サイズの数だけ、設定したパラメータを持つワイブル分布に従う乱数を発生させて、これを元の直径データ（倍精度実数で、単位は cm）とした。この値から、(1) 倍精度実数のまま、(2) 0.1cm 単位で丸めた値、(3) 2cm 括約の直径階分布、(4) 5cm 括約の直径階分布、の 4 種類の精度のデータを生成し、それぞれについてワイブル分布のパラメータを、マルコフ連鎖モンテカルロ (MCMC) 法によりベイズ推定した。シミュレーションには、統計ソフトウェアの R 3.6.3 (R Core Team 2020) および Stan 2.25.0 (Stan Development Team 2020) を使用した (List S1–3)。なお、R と Stan は両者とも、実数の実装は IEEE 754 規格に準拠しており、10 進数では 15 桁の精度がある (R Core Team 2020, Stan Development Team 2020)。シミュレーションでは、データが (1) と (2) の場合は、それぞれの直径の値をそのままデータとして与えて (List S2)、(3) と (4) の場合は、胸高直径階ごとの本数をデータとして与えて (List S3)、各パラメータの事後分布を推定し、事後平均値を求めた。このシミュレーションを、6 通りの組み合わせについてそれぞれ 1000 回繰り返し、事後平均値の平均と標準偏差を求め、データの精度による違いを比較した。

結果を Table 1 に示す。標本サイズ 300、形状パラメータ 2、尺度パラメータ 8、精度 0.1cm 場合に 10 回だけ Stan による MCMC 計算が失敗するケースがあったが、これは生成されたデータに対して MCMC 計算の初期値が適合しないためであった。この場合を除いて、パラメータの推定値の平均と標準偏差を求めた。データの精度が、0.1cm 単位、2cm 括約、5cm 括約となっても、倍精度実数の場合と事後平均値は大きくは変わらなかった。ただ、標本サイズが 30 の場合で形状および尺度パラメータの真値が 8 以下の場合には、データの精度にかかわらず、事後平均値は真値よりも多少大きくなる傾向があった。これは、小標本の場合には推定値のばらつきが大きくなるため、とくに真値が小さい場合には、正值のみをとるワイブル分布のパラメータ推定値の事後分布が右に裾が広がったためであろう。このほか小標本では、とくに形状パラメータにおいて 5cm 括約では標準偏差が大きくなり、推定値の精度が悪化する傾向が見られた。また尺度パラメータでは、真値が大きくなるほど標準偏差が小さくなる傾向がいずれの標本サイズでも見られたが、この理由としては、同時に形状パラメータを大きくしているため、その影響を受けたことが考えられる。

以上の結果から、シミュレーションで確認した範囲のパラメータにおいては、とくに大標本であれば、5cm 括

約ほどに精度を落とした胸高直径階のデータを用いても、ワイブル分布のパラメータをおおむね正確に推定できると考えられる。

本報告は、森林総合研究所交付金プロジェクト「天然更新による低コストカンバ施業システムの開発」(課題番号 201903) の成果である。また、森林総合研究所北海道支所の中西敦史博士には、本報告の草稿を読んでいただき、有用なコメントをいただいた。お礼申し上げる。

引用文献

- Bailey, R. L. and Dell, T. R. (1973) Quantifying diameter distributions with the Weibull function. *For. Sci.*, 19, 97–104.
- Cheng, C.-P., Wei, C., Tsai, M.-J. and Tsao, T.-M. (2017) A Weibull model of the impact of thinning and a typhoon event on the stand structure of *Cryptomeria japonica* in Central Taiwan over 100 years. *J. For. Res.*, 22, 22–29.
- 大野 泰之 (2018) シラカンバ人工林の生育実態—径級別の原木供給ポテンシャルの試算— *北方林業*, 69, 97 – 103.
- Rennolls, K., Geary, D. N. and Rollinson, T. J. D. (1985) Characterizing diameter distributions by the use of the Weibull distribution. *Forestry*, 58, 57–66.
- R Core Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/> (参照 2021-02-24).
- Stan Development Team (2020) Stan reference manual version 2.25. https://mc-stan.org/docs/2_25/reference-manual/index.html (参照 2021-02-24).

補足電子資料

以下はオンライン版のみの掲載となります。

<https://www.ffpri.affrc.go.jp/pubs/bulletin/461/index.html>

List S1. シミュレーションを実行する R スクリプト

List S2. 直径データをワイブル分布にあてはめてパラメータを推定する Stan のコード

List S3. 直径階分布データをワイブル分布にあてはめてパラメータを推定する Stan のコード