

令和元年度

国有林における木材供給手法に係る
立木価格調査検討業務報告書

令和2年2月

林 野 庁

目 次

1. 事業の概要-----	1
(1) 目的-----	1
(2) 事業の内容-----	1
2. 事業の実施体制-----	2
(1) 実施主体-----	2
(2) 実施体制-----	2
(3) 検討委員会の構成-----	2
(4) 検討委員会の開催と検討事項-----	2
3. 事業実施に当たっての基本的な考え方-----	3
(1) 樹木料の算定-----	3
(2) 樹木料評定額の算定方法の検討-----	4
4. 分析-----	6
(1) 目的-----	6
(2) 方法-----	6
(3) データ-----	11
(4) 推定結果-----	22
5. 検証-----	32
(1) 検証に使用したデータ-----	32
(2) 検証結果-----	32
6. まとめ-----	39
7. 参考資料	

1. 事業の概要

(1) 目的

令和元年6月5日に可決、成立した「国有林野の管理経営に関する法律等の一部を改正する法律」に基づき、国有林において樹木採取権制度が創設されることになった。

樹木採取権制度では、国が樹木採取権者から樹木の対価として樹木料を徴収することとしている。本事業においては、地域や林分状況等に応じた適切な樹木料を算定するために、主に立木販売の実績データ等を用いた統計分析を行い、樹木料算定に当たっての基礎資料を作成することを目的とする。

(2) 事業の内容

本事業は、「令和元年度国有林における木材供給手法に係る立木価格調査検討業務委託事業仕様書」に基づき実施した。ただし、具体的に分析するデータ及び分析手法は検討委員会における議論を踏まえ決定された。例えば、当初は、販売額を予定価格で除した「値開率」を用いた算定方法を想定しており、値開率の分析を主眼としていたが、検討委員会において、値開率による算定方法は統計学的に適切ではないことが指摘された（4. 分析を参照）ことを踏まえ、販売額を被説明変数とした重回帰式の作成が主たる目的となった。実際に分析に用いたデータ、分析手法及び検証作業の内容は下記のとおりである。

①分析に用いたデータの収集・分析

分析に必要なデータとして、立木販売実績における販売額、予定価格及びその算定因子である収穫調査データ（場所、時期、林分条件、材積等）等を収集した。

収集に当たっての主な観点として、地域による影響（地域ごとの需給動向等）、時期による影響（時期ごとの市況や物価指数等）、樹種や材積等の林分状況、地形や林道距離等の搬出条件による影響、伐採方法（主伐、間伐）の違い等が挙げられる。本事業において収集及び分析したデータの詳細については4.分析に記す。

②立木販売の販売額を予測する回帰式の作成

検討委員会での討議を踏まえ、樹木料の算定方法の基礎資料として、国有林の過去の立木販売データから、販売額を予測する回帰式を作成した。

また、回帰式の精度について、実際の立木販売の実績との比較により検証を行った。

③運用ルール案の作成

実際に樹木料を算定するに当たって、エクセルによる算定様式案を作成した。

2. 事業の実施体制

(1) 実施主体

一般財団法人日本木材総合情報センターに委託して実施した。

(2) 実施体制

事業の実施体制は下記のとおりである。

氏名	所属	役割分担
井上 幹博	業務執行理事	統括責任者
永井 寛	国内情報部長	事業責任者
平沢 敏正	総務部長	契約・経理担当者
武田 八郎	国内情報部	既存統計データ収集・分析
上島 信彦	国内情報部	国有林データ収集・加工・データ解析
佐々木 央	国内情報部	資料作成・事務
藤掛 一郎	宮崎大学	樹木料評定額の算出式の作成、検証

(3) 検討委員会の構成

本事業の実施に当たっては、木材市況や統計分析等に造詣のある学識経験者から構成された「国有林における木材供給手法に係る立木価格調査検討委員会」を設置し、主にデータの収集、統計分析の手法の決定、回帰式の作成及び検証に当たって指導、助言を得た。

検討委員会の構成は下記のとおりである。

氏名	所属・役職
遠藤 日雄※	特定非営利活動法人活木活木森ネットワーク 理事長
久保山裕史	国立研究開発法人森林研究・整備機構森林総合研究所林業経営・政策領域長
立花 敏	国立大学法人筑波大学生命環境系 准教授
藤掛 一郎	国立大学法人宮崎大学農学部森林緑地環境科学科 教授
吉本 敦	大学共同機関法人情報・システム研究機構統計数理研究所 教授

※座長

(4) 検討委員会の開催と検討事項

検討委員会は計4回開催した。検討事項は下記のとおりである。

第1回検討委員会

日時：令和元年10月2日 13:30～16:30

場所：林友ビル6階会議室

議事

- (1) 立木価格調査検討業務の実施目的と内容について
- (2) 値開率の算定方法（案）について

第2回検討委員会

日時：令和元年11月1日 13:30～16:30

場所：林友ビル6階会議室

議事

- (1) 値開率についての分析結果
- (2) 算出のための Excel ブックのイメージ
- (3) 過去5年間の木材価格・需給の分析

第3回検討委員会

日時：令和元年12月9日 13:30～16:30

場所：林友ビル6階会議室

議事

- (1) 市場価格モデル案の運用に向けた検討について
- (2) 市場価格算定モデル案について

第4回検討委員会

日時：令和2年1月28日 13:00～16:00

場所：日中友好会館地下 第1会議室

議事

- (1) 樹木料評定式について

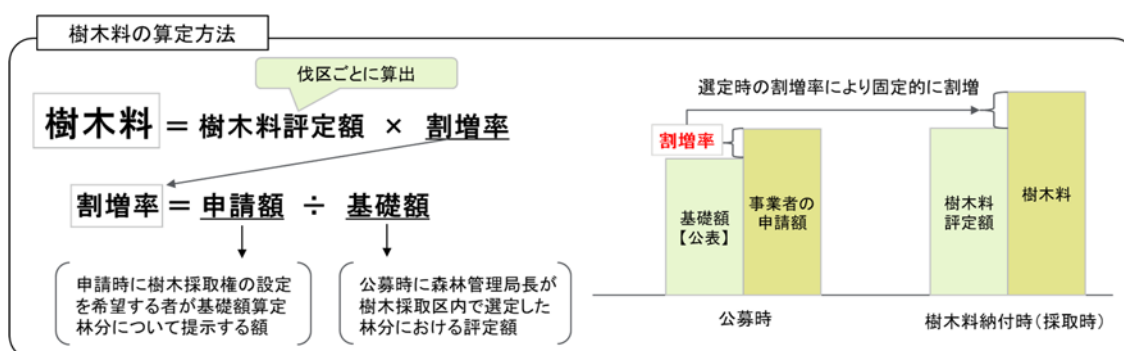
3. 事業実施に当たっての基本的な考え方

(1) 樹木料の算定

従来行われてきた一般競争入札による立木販売では、国有林は契約物件ごとに予定価格（非公表）を算出し、予定価格以上の金額で最も高い入札価格を提示した者にその額で販売するという方法で値決めを行ってきた（ここで予定価格や販売額はいわゆる単価（ m^3 当たりの価格）ではなく、個々の契約物件に対して付けられる総額である）。

これに対し、樹木採取権制度においては、公募時に国が選定した林分及びその林分の国としての評定額（以下、「基礎額」という。）を公表し、これに対し、申請者は当該林分について自らが見積もった額（以下、「申請額」という。）を提示する。申請額が基礎額を上回る率を割増率とし、樹木採取権の存続期間内に順次行われる採取に際しては、その時々、国が公募時に示した方法で算出した額（以下「樹木料評定額」という。）に割増率を乗じて、個別の採取対象の林分の樹木料を算定することとしている。

このように、従来の立木販売と異なる価額の決定方法であることから、樹木料の決定に当たっては透明性、公平性が確保されている必要があり、そのため、公募時に、樹木採取権者にその算定過程がわかるように公表するものである。



本事業の目的である樹木料の算定に当たっての基礎資料の作成とは、「樹木料評定額」の算定方法案の作成であり、樹木料評定額は、「4. 分析」に後述するが、国有林における立木販売における販売額相当の価額を想定している。

なお、樹木料の具体的な運用に関しては「樹木採取権制度ガイドライン（令和2年3月公表予定）」に詳述することとし、本稿では割愛する。

（2）樹木料評定額の算定方法の検討

①立木販売と樹木採取権の価格決定プロセスの違い

現行の立木販売は、通常は一般競争入札により実施されている。一般競争入札では、国が決定する最低額である予定価格（非公表）以上の金額で、最も高い金額で応札した者が落札し、いずれの応札金額も予定価格を下回った場合は不落札となる。このとき、国が決定する予定価格は、市場価逆算方式により算定されている。

樹木採取権における樹木料は、対象となる樹木の対価である。そのため、樹木料は、立木販売において、国が市場価逆算方式により決定する最低額としての予定価格とは異なり、一般競争入札により値開率を含めた価額の算定方法の検討が必要となる。

②統計分析手法による客観性のある算定方法

樹木料評定額を求めるにあたって、算定方法に客観性を持たせるため、統計的手法を用いることとした。統計的手法は、多変量解析の中で最も一般的で、社会科学分野でも多く用いられている重回帰分析が適当であると考えた。重回帰分析は結果となる被説明変数に対して、原因となる複数の説明変数がどのような影響を与えているのか明らかにすることができるからである。

樹木料評定額の算定には重回帰分析を適用することは検討委員会においても妥当であると判断された。このことから、本事業の具体的な目的は、樹木料評定額を求めるための重回帰式（以下「樹木料評定式」という。）を作成することとなった。

なお、重回帰分析における被説明変数及び説明変数については、検討委員会の議論を踏まえて決定され、第2回検討委員会までの議論を踏まえ、値開率を求める方法が統計学的に適切ではないことが指摘されたことを踏まえ、被説明変数は立木販売の販売額とした。

③樹木料評定式の更新の考え方

樹木採取権の存続期間は10年を基本（最長50年）に運用することとしており、樹木料評定式は、樹木採取権の公募の際に公表し、権利の存続期間に渡って樹木料評定式は基本的に変えないこととしている。

10年間、同じ樹木料評定式を適用することの是非については検討委員会においても議論となったが、樹木料評定式の変数には、近隣の原木市場等の丸太単価や木材生産に係る経費など、樹木料評定時点の状況を反映している変数が含まれており、これらの変数の変動により時々の状況が適切に反映されるものと考えられるためである。

また、公募時において樹木料評定式は公表され、その内容を踏まえて樹木採取権が設定されるため、特に樹木採取権の存続期間中に必要以上に樹木料評定式を変更することは、かえって公平性、公正性を欠く結果となることも考えられる。

一方で、樹木料評定式の変数だけではとらえきれないような極めて著しい経済その他の状況変化等が認められた場合等も想定されるため、検討委員会においても、PDCAサイクルにより樹木料評定式の妥当性について検証することが必要であるとされた。

4. 分析

(1) 目的

本調査の目的は樹木採取権制度における樹木料評定額の算定方法の検討のための考え方を整理し、基礎的な資料を作成することである。本節ではそのために、統計手法を用い、立木販売等のデータから販売額を予測する式の作成を試みる。

新たな制度における樹木料評定額も従来の立木販売における予定価格も、対象となる樹木の評価額であることに変わりはない。しかし、いわゆる指値を定める役割を持つ従来の予定価格の設定とは異なり、樹木料評定額は従来の立木販売における販売額に相当する。また、割増率は、買い手側の物件への標準的な評価を国が定め、これを超える分を明らかにしておくという役割がある。樹木採取権制度では、従来は個別に立木販売していた物件をまとめ、それを長期に渡って順次採取（＝伐採し取得）していく権利を設定する。買い手にとっては、複数物件を一括して確保できるため、量がまとまり、立木調達が長期にわたり安定するメリットがあり、さらに、近接する物件をまとめて確保できることで伐出を一体的に行うなどして作業効率を高められるといったメリットも考えられる。これらのメリットの分、樹木採取権制度は個々の物件が個別に立木販売をするより高い価値を持ちうることを想定している。そこで、この販売額の増加分を割増率として算定しておこうというわけである。そのため、樹木料評定額は、樹木採取権設定の対象となる立木販売物件が、従来どおり個別物件ごとに一般競争入札で販売されればいくらの販売額で売れたであろうかという意味での評価額でなければならない。この意味で、樹木料評定額は従来の販売方法で言えば、予定価格ではなく、販売額に相当するものと考えられる。

以上のことから、本節での分析の目的は、主に従来の一般競争入札による立木販売の結果をデータとし、各物件を構成する種々の情報からその物件の販売額がいくらになるかを予測する式を作成することである。

(2) 方法

立木販売物件の販売額を、その物件の各種情報から予測する式が必要である。このような予測式の作成には、多数の変数間の関係を分析する多変量解析の方法の中でも重回帰分析が適しており、これを採用した。重回帰分析は、複数の説明変数から、一つの被説明変数の値を予測する式を実際のデータから推定する方法である。例えば、立木販売物件の販売額を S とし、各種の物件情報を X_1 、 X_2 、...、 X_k とする時、

$$S = \alpha_0 + \alpha_1 \times X_1 + \alpha_2 \times X_2 + \dots + \alpha_k \times X_k + \text{誤差}$$

といった関係が成り立つものと想定する。この販売額 S と各種物件情報 X_1 、 X_2 、...、 X_k との関係式を、回帰式と呼ぶ。回帰式の係数 α_0 、 α_1 、 α_2 、...、 α_k を特定することができれば、各種物件情報 X_1 、 X_2 、...、 X_k から販売額 S を予測することができるようになる。

そこで、重回帰分析では、 S 、 X_1 、 X_2 、 \dots 、 X_k に関する n 個のデータ、

$$(S_1, X_{11}, X_{21}, \dots, X_{k1})$$

$$(S_2, X_{12}, X_{22}, \dots, X_{k2})$$

...

$$(S_n, X_{1n}, X_{2n}, \dots, X_{kn})$$

を使って、係数 α_0 、 α_1 、 α_2 、 \dots 、 α_k の最も確からしい値を決める（推定する）ことで、 X_1 、 X_2 、 \dots 、 X_k から S を予測する式を得る。つまり、誤差の部分を省くと、 S と X_1 、 X_2 、 \dots 、 X_k の関係は、

$$S = 4 + 2X_1 - 5X_2 + \dots + 3X_k$$

という関係がもっともらしいのか、それとも

$$S = 2 - 3X_1 + 8X_2 + \dots + X_k$$

という関係がもっともらしいのか、あるいは他のどんな係数の組み合わせがもっともらしいのかを実際に観測された S 、 X_1 、 X_2 、 \dots 、 X_k に関する n 個のデータから導き出すのである。

従って、重回帰分析では、（1） S を予測する式にはどの変数を説明変数として使えばよいかを特定することと、（2）それぞれの説明変数の係数の値は何がよいかを決めることを行う。このうち、まず（2）の係数を決める方法として、重回帰分析で用いられる最も標準的な方法に、通常の最小二乗法（Ordinary Least Squares、略して OLS）という方法がある。この方法を用いると、係数の推定結果が望ましいものとなるという各種の知見が得られており、それゆえに重宝されている。本分析でも、この OLS 回帰による係数推定を行う。

ただし、OLS 回帰は唯一の係数推定法ではなく、他にもそれぞれ特徴を持った係数推定法がある。その中で本分析では、中央値回帰という方法をもう一つの方法として採用する。詳細の説明は省くが、重回帰分析では、被説明変数である S の確率分布を考える。そして、OLS 回帰は、説明変数 X_1 、 X_2 、 \dots 、 X_k が定まった時の S の確率分布の平均を予測しようとする手法である。例えば、 X_1 が 3、 X_2 が -10、 X_k が 7 と決まった時に、 S の値は必ずしも一つには決まらず、ばらつきを持つ（確率分布する）だろうが、その分布の平均（中心）がどこに位置するかは X_1 が 3、 X_2 が -10、 X_k が 7 であることに影響を受けるだろう。その関係を、 X_1 、 X_2 、 \dots 、 X_k にかかる係数の値を決めて、 $2 - 3X_1 + 8X_2 + \dots + X_k$ のような式で言い当てようとするのである。これに対して、中央値回帰による係数推定では、説明変数が定まった時の S の確率分布の平均ではなく中央値を言い当てる式を特定しようとする。中央値とは、変数がばらつきを持って分布する中の、下から 50% 目の値のことである。

もし、 S の確率分布が左右対称であれば、その分布の平均と中央値は一致する。そのため、OLS 回帰を用いて平均を言い当てようとする式を作っても、中央値回帰を用いて中央値を言い当てようとする式を作っても、同じ式が得られることが期待され、その意味では、どちらか一方の方法を用いれば十分である。

しかし、今回の分析で被説明変数として取り上げる立木販売物件の販売額の分布は左右対象の分布ではない。むしろ分布としては右側の裾の長い分布が想定される。というのも、販売額は 0 円を下回ることはもちろんないし、それだけでなく、予定価格を下回ることもない。応札価格が全て予定価格を下回った場合にはその入札は不調に終わり、販売がなされないからである。そのため、通常、販売額は予定価格より少し高いところに集中して分布するが、時には予定価格より大幅に高い価格で販売されることもある。こうして販売額の分布は左右非対称で、右裾の長い分布であることが想定される。

このような分布の場合、平均と中央値は異なる値を取る。具体的には、平均が中央値より大きな値を取る。平均と中央値はどちらもそれぞれの意味において分布の中心を示す量であり、今回の販売額の予測においても、どちらを予測するのが正しいとあらかじめ判断できるわけではない。そこで、まず OLS 回帰と中央値回帰の両方を行う。そして、それぞれで特定された式を使って、販売額の予測がどれだけうまくいくのかを比べ、OLS 回帰で得られた予測式と中央値回帰で得られた予測式のどちらを使うかを選択することとする。

次に、(2) の販売額を予測するための説明変数として何を使うべきかという問題については、OLS 回帰の場合も中央値回帰の場合も、Stepwise 変数減少法を使った変数選択を行うこととした。その際の変数選択の基準としては主に係数の有意性を表す P 値を用いた。ここでの P 値とは、ある変数 X_i の係数 α_i の真の値が 0 である確率を評価した値で、0 から 1 の間の値を取る。P 値が小さく 0 に近いほど、係数 α_i が 0 であるとは考えにくい、つまり、変数 X_i が被説明変数 S に何らかの影響を与えている可能性が高いことを意味する。このことを、変数もしくはその係数が統計的に有意であるという。

本分析で採用した Stepwise 法では、まず初めに説明変数の候補と考えられる変数を全て使い回帰式を推定する。その推定結果において、最も P 値が大きく、従って有意性が低い変数を除外し、変数を一つ減らして回帰式を推定し直す。そして、再び推定結果において最も P 値が大きく有意性が低い変数を除外する。これを繰り返していき、全ての残った変数の P 値が一定の基準値を下回ったところで、残っている説明変数を採用するという方法を取る。従って、 S に有意な影響を与えていない変数を除いていき、一定の有意性を持つ変数だけで予測式を構成することになる。本分析では、変数を除去しないこととする P 値の基準値を 0.2 とした。基準として 0.1~0.2 が使われることが多いが、今回はモデルの予測精度を重視し変数が残りやすいことを考慮して 0.2 とした。

加えて、OLS 回帰については、Stepwise 法の基準に P 値ではなく AIC を用いる方法も併用した。AIC は回帰式全体がいかに少ない変数で当てはまりのよい式となっているかを評価する指標である。具体的に、AIC は値が小さいほど、回帰式全体が効率的で当てはまりがよいことを

表す。そこで、AIC を用いた Stepwise 変数減少法では、全ての説明変数を使うところからスタートし、変数を落とすことで AIC を最も大きく下げうる変数一つずつ落としていき、それ以上、AIC が下がらなくなったところで、残った説明変数を採用する。しかし、中央値回帰は OLS 回帰と異なり尤度に基づく方法ではないことから AIC は利用できない。そのため、OLS 回帰についてのみこの方法を併用した。

続いて、被説明変数を販売額とすることに関しても説明を加えておこう。既に触れたように、販売額は、物件に対する最高入札価格が予定価格を上回り不落にならなかった場合だけ得られる限定されたデータである。このことを、販売額は最高入札価格の切断データであると言う。切断データとは、ある変数の分布の中で、その変数が一定の値を下回ったり、上回ったりした部分が観測できず、欠落したデータのことである。重回帰分析の手法の中には、いくつかの仮定を置きながら、切断されたデータを用いて、切断される前の分布が持つ平均を予測する方法もある。しかし、新たに始まる樹木採取権制度の対象には、立木販売において不落札になるような物件は含まれないであろうと想定されており、今回の分析では、予定価格を下回った場合も含めた最高入札価格の分布、また、その平均や中央値が知りたいわけではなく、順調に予定価格を上回る入札があり、落札された場合の販売額の平均や中央値を予測することが目的である。そのため、入札が成立した立木販売物件のデータだけを使い、販売額の予測式を作成することとした。

また、販売額に代わる被説明変数として考えられるものとして、値開率がある。値開率は次の式で定義される。

値開率 = 販売額 / 予定価格

値開率は販売額が予定価格の何倍であったかを示し、入札が成立し販売が行われたデータでは必ず 1 以上の値を取る。値開率はこのように予定価格と販売額との関係を表す指標として使われてきたもので、当初この値開率を被説明変数とすることも検討した。しかし、値開率を予測する式を得ることは可能だが、最終的に得たいのは販売額の予測値であるから、値開率を予測した場合、予定価格に値開率の予測値を乗じて、販売額を予測することになる。しかしそれならば、初めから販売額を予測する式を推定の方がよいと判断した。このことについて説明する前に、まずは OLS 回帰や中央値回帰がどのようにして予測式の係数を特定するかについて説明しておく。

OLS 回帰は、説明変数が特定の値を取った時の被説明変数の平均を予測する式を作るもので、中央値回帰は、被説明変数の中央値を予測する式を作るものであると先に説明した。これらの回帰分析の手法を、また別の角度から特徴付けることができる。すなわち、OLS 回帰は、被説明変数の予測値と実現値の差（これを残差と呼ぶ）を取り 2 乗したものについて、標本全体での合計が最小になるように各変数の係数を決める方法でもある。また、中央値回帰は、被

説明変数の予測値と実現値との間の残差の絶対値を標本全体で合計したものが最小になるように各変数の係数を決める方法である。

例えば、係数をある一つの組み合わせに決め、 $4 + 2X_1 - 5X_2 + \dots + 3X_k$ の予測式を決めたとする。データとして、 S 、 X_1 、 X_2 、 \dots 、 X_k に関する n 個のデータ、 $(S_j, X_{1j}, X_{2j}, \dots, X_{kj})$ ($j=1,2,\dots,n$) が得られているので、その一つ一つについて、 X_{1j} 、 X_{2j} 、 \dots 、 X_{kj} の値を $4 + 2X_1 - 5X_2 + \dots + 3X_k$ の予測式に代入すれば、販売額の一つの予測値 S_j^{\wedge} (「 \wedge 」は予測値であることを示すものとする) が得られる。この S_j^{\wedge} と実際に観測された販売額の実現値 S_j とから、残差平方和を作る。

$$\text{残差平方和} = \sum_j (S_j^{\wedge} - S_j)^2$$

この残差平方和は係数の組み合わせを変えれば、予測値 S_j^{\wedge} を通じて値が変わる。そこで、OLS 回帰は、この残差平方和が最も小さくなるように係数の組み合わせを決める方法である。また、中央値回帰は、同様の表記を使えば、

$$\text{絶対残差和} = \sum_j |S_j^{\wedge} - S_j|$$

を最小にするように係数の組み合わせを決める方法である。

このようにこれら二つの方法は、残差の平方を使うのか、残差の絶対値を使うのかの違いはあるものの、全体として予測結果と実際の観測結果との差を最小にするように予測式を特定する方法である（平方または絶対値を使うのは単なる残差の和ではプラスとマイナスの残差が打ち消し合い全体としての誤差の大きさを表さないためである）。

そうであれば、販売額を精度高く予測することが目的の場合には、販売額を被説明変数として販売額の予測式を得るのが、一般的にはよい方法である。値開率を被説明変数とした場合には、値開率を精度高く予測できる式は得られるが、それに予定価格を乗じて販売額を得た場合には、販売額の予測精度が高いという保証はなく、販売額を被説明変数とする予測式からの予測精度を上回ることがあり得ないわけではないが、考えにくい。実際、値開率の予測式を推定し、値開率の予測値に予定価格を乗じて販売額の予測値を得る方法も試したが、やはりダイレクトに販売額を予測した方が残差が小さい傾向が認められたため、販売額を被説明変数とした。

以上のことから、立木販売物件の販売額を被説明変数とし、物件の各種情報を説明変数とする回帰式を作り、OLS 回帰、中央値回帰によって係数を推定して、販売額の予測式を得ることとした。そして、最終的に回帰式で使用する変数は Stepwise 法によって選択するものとした。

最後に、推定に使用する立木販売物件データは全国 7 つの森林管理局にまたがるものであった。従って、全国のデータで一つの回帰式を推定することも可能だが、今回の分析では、北海道森林管理局（以下「北海道局」という。）とそれ以外の森林管理局（以下、「本州以南の

局」という。)との2つに分けて、それぞれのデータで回帰式を推定し、北海道局用の販売額予測式と本州以南の局用の販売額予測式を得ることとした。この理由として、1点目は、北海道局と本州以南の局では、後に詳しく述べるが、得られる変数に若干の違いがあったことが挙げられる。それらの変数を使わなければ、一つの式で全国をまとめることができるが、それらの変数の中には販売額に有意に影響を与えているものがあつた。2点目として、後に結果を提示する際に説明するが、北海道と本州以南では、立木販売物件の内容にも差があつた。北海道では間伐の物件が多く、本州以南では皆伐の物件がほとんどであつた。3点目として、北海道局は立木販売実績数が多く、北海道局の実績だけで統計分析をするに当たって十分なデータを得ることができたことが挙げられる。これらのデータの特徴による理由に加えて、実際の運用面からみても、地形、地質等が北海道と本州以南で異なっていることから、北海道と本州以南でそれぞれ別に式を推定することが妥当であると考えられた。

(3) データ

回帰式の推定に用いるデータとして、2014(H26)～18(H30)年度の過去5年分の全国の立木販売物件のデータを準備した。ただし、この中には樹木採取区の対象にはなり得ない物件が存在する。そのため、その中から、まず①販売方法が一般競争入札以外の物件は除外した。すなわち、随意契約などの方法で販売された物件は除外した。元のデータは契約物件数で15,425件であつたが、随意契約等を除外することで、対象は5,150件に絞られた。加えて、イレギュラーな物件を除外するために、②混合契約、③林名区分が官行造林である物件、④施業方法が育成単層林と育成複層林のいずれか以外の物件(=天然生林)、は除外した。②は立木販売と造林事業との混合契約である。③の官行造林は、国有林野ではないため除外することとし、残った区分として、分収林を除いた国有林(以下、便宜的に「直営林」という。)、分収林である分収造林及び分収育林の3つの立木販売物件が分析の対象となる。これら②～④の限定により、対象物件は4,546件となった。さらに、説明変数候補のいずれかのデータに欠落があるもの81件、風倒被害木を含むなど特殊事情のあるもの84件を除外して、4,381件が推定に使用したデータであつた。

表1はその4,381件を、本州以南か北海道の別、また直営林、分収育林、分収造林の別の内訳を見たものである。本州以南の局と北海道局では、物件数がそれぞれ3,411件と970件であつた。北海道局ではほぼ80%が直営林での販売であるのに対して、本州以南の局では55%が分収造林、20%が分収育林で、直営林は24%に過ぎなかつた。樹木採取権制度での立木販売は直営林の販売が想定されているため、データを直営林に限定することも考えられたが、それでは本州以南の物件数が大幅に減ってしまう。分収林であれ直営林であれ、販売額と各種物件情報との関係に違いはないものと考え、分収育林、分収造林のデータも使用することとした。ただし、説明変数に分収育林と分収造林のダミー変数を加え、区分の違いが販売額に与える影響がもしもあるとすれば、ダミー変数の効果として吸収できるようにした。

表2は、やはり推定に使用した4,381件の販売年度別の件数である。2014年度は少ないが、各年度に分散している。

表1 推定に使用した物件の地域・林名区分別内訳

	本州以南		北海道		計	
	件	%	件	%	件	%
直営林	830	24.3	783	80.7	1,613	36.8
分収育林	692	20.3	103	10.6	795	18.2
分収造林	1,889	55.4	84	8.7	1,973	45.0
計	3,411	100.0	970	100.0	4,381	100.0

表2 推定に使用した物件の販売年度別内訳

契約年度	本州以南	北海道	計
2014年度	473	162	635
2015年度	685	218	903
2016年度	700	209	909
2017年度	780	190	970
2018年度	773	191	964
計	3,411	970	4,381

なお、これらの立木販売実績に係るデータは「国有林野情報管理システム」から抽出したものを使用している。

また、立木販売における販売額は契約単位で決定されるが、1件の契約には複数の林小班が含まれることがあり、面積や（結果的に使用しなかったが）森林調査簿のデータ等、林小班毎のデータは契約毎のデータに変換を行っている。同様に、本数や材積、林齢等のデータは樹材種毎のデータであり、これらも契約毎のデータに変換して分析を行っている。

実際に推定に使用した変数について、まず、被説明変数は既に説明した通り、販売額とした。樹木料評定額は税抜額を想定しているため、販売額は税抜販売額を用いた。そして、説明変数には、各種物件情報を変数化した次の①～⑥のグループのものをを用いた。順に説明する。

①丸太価額、生産変動経費、生産固定経費（及びこれらの単価）

立木販売に当たっては販売物件の予定価格が設定され、販売額は予定価格を上回る額となる。そのため、予定価格は販売額を予測する上で重要な要因である。この予定価格はいわゆる市場価逆算法によって立木の価値を推計することで求められる。その計算に使われるのが、丸太価額、生産変動経費、生産固定経費である。丸太価額は、収穫調査の結果及び市場調査の結果をもとに、対象の林分から生産が見込まれる丸太の規格別の材積（m³）に、近隣の原木市場等の丸太価格（円/m³）を乗じたものである。生産変動経費は生産に必要な経費をやはり収穫調

査の結果などから算出したもので、生産変動経費は生産が見込まれる丸太材積に応じて変化する経費であり、伐木造材、運搬、販売に要する事業費である。生産固定経費は搬出施設費や小屋掛費等で材積按分を適当としない固定的な経費部分である。そこで、本分析では販売額を規定する要因として、まずはこれら丸太価額、生産変動経費、生産固定経費を用いることとした。予定価格を構成する関係から、販売額に対し丸太価額は正（プラス）、生産変動経費と生産固定経費は負（マイナス）の影響を持つことが予想される。しかし、これらの因子はそもそも予定価格算定において使用されており、今回のように販売額を求める際に必ずしもそのような符号条件が満たされるとは限らない。

さて次に、その他の説明変数として何を採用するかであるが、そもそも、販売額は予定価格との関係では、

$$\text{販売額} = \text{予定価格} + (\text{販売額} - \text{予定価格}) = \text{予定価格} + \text{予定価格と販売額との差分}$$

と分けて捉えることができ、丸太価額、生産変動経費、生産固定経費は、この右辺第1項の予定価格部分を通じて主に販売額を規定すると考えられる。すると次には、残る「予定価格と販売額との差分」を説明するのはどのような変数であるかが問題になる。国有林が一定の方式で市場価逆算法で算定した予定価格と実際に入札で表明された買い手の最高入札価格で示された評価との差はどのような要因によるものであろうか。しかし、残念ながら、このことを予め明らかにしうる知見はない。

そこで、以下、販売物件に関する各種の情報を可能な限り変数化して説明変数候補とすることで、「予定価格と販売額との差分」の部分を含めた販売額を予測する式を作成することとした。例えば、もしかすると、面積が大きな物件では、「予定価格と販売額との差分」が大きい、あるいは小さいといった関係があるかもしれない。あるいは、林齢が高い物件かどうかそこに影響しているかもしれない。様々な可能性が考えられ、また、どれが確からしいかについて、頼りになる情報はない。そこで、こうした変数をできるだけ多く準備して説明変数の候補とすることにした。

こうした変数の候補として、まず、丸太価額、生産変動経費、生産固定経費をそれぞれ丸太材積で除した、1 m³当たりの丸太単価、生産変動経費単価、生産固定経費単価を取り上げることにした。丸太単価が高い物件は高価な材を多く含む物件である。また、生産固定経費には輸送費が含まれるため、生産固定経費単価が高い物件には地利条件が悪い物件が含まれる。こうした条件の違いが販売額に影響している可能性が考えられる。

②面積、立木材積、単位蓄積、単木材積、本数密度、樹種比率（スギ、ヒノキ、トドエゾ、カラマツ、その他針葉樹、広葉樹）、林齢、平均胸高直径、平均樹高

次に、販売物件の林分内容を計測した結果得られる変数群である。収穫調査の結果、一つの販売物件の内容が、契約に含まれる林小班、樹材種ごとに面積や立木材積などが記された複数

行のデータにまとめられる。まずはこれを合計して、販売物件の面積、立木材積が得られる。また、同様に合計の立木本数も分かるので、それらから、単位蓄積（＝面積当たりの立木材積）、単木材積（＝本数当たりの立木材積）、本数密度（＝面積当たりの立木本数）を得た。なお、物件の規模については、面積で見て 1ha 未満のものから 100ha を超えるものまで 4 桁の差があり開きが大きいので、規模の影響が線形でない可能性も考えられることから、折れ線回帰を検討することとした。最終的には、物件規模と C 経費との関係などを検討の上、回帰式の当てはまりの良さも考慮し、面積が 10 ha 以上で折れ線となるよう、面積が 10 ha までは 0、10ha を超えると面積-10 を取る変数「面積 10ha 超」を作り、面積の変数と合わせて用いることとした。

一つの販売物件には、複数の樹種、異なる林齢の立木が混ざっていることが多い。そこで、樹種比率や林齢を次のように変数化した。樹種比率は、詳細に分類されている収穫調査の樹種をスギ、ヒノキ、トドエゾ（トドマツ及びエゾマツ）、カラマツ、その他針葉樹、広葉樹の 6 つに分類し、それぞれの比率を算出したものである。

林齢は、物件内で最も立木材積の多い樹種の林齢をその物件の林齢として使用した（樹種は上記のスギ、ヒノキ、トドエゾ、カラマツ、その他針葉樹、広葉樹の 6 つにまとめたもの）。樹種毎の材積に差がないときは、林齢の低い方を使った。また、同じ樹種でも林齢が複数ある時は、立木材積の多い方の林齢を使った。ただし、おそらく欠測だと思われるが、データ上で林齢が 0 となっている林分や、あるいは多様な林相からなる物件の場合に一部に林齢が 1 となっている林分が含まれていることもあり、これらは除外した上で、上記ロジックを適用した。

契約に含まれる林小班、樹材種ごとに本数や面積のほか、林齢や胸高直径が記載されているので、これから契約に含まれる全林木の 1 本当たりの平均胸高直径と平均樹高を算出した。ただし、平均胸高直径と平均樹高は、国有林野情報管理システムにおいて通常整理はされていないため、北海道局ではこれらのデータは得られず、本州以南の回帰式推定においてのみ用いた。

③立木材積における低質材材積比率、丸太材積における一般材、低質材、原料材材積比率

収穫調査では、材種が記録される。材種は一般材や低質材などの別を示したものであるが、北海道局と本州以南の局では記録の仕方が異なっており、今回分析に使用したデータにおいて、北海道局では全て一般用材として整理されていた。北海道局では、それとは別に、対象の林分から生産が見込まれる丸太材積における A 材（一般材及び特殊用途材）、B 材（低質材及び原料材）、C 材（幹材の生産に伴って発生する末木及び枝条から生産される原料材（長級 1.8 m 未満を含む。））の別ごとのデータが整理されており、分析に利用可能である。そこで、本州以南については、材種を用いて、低質材率を全体の立木材積に占める低質材の割合として算出し、北海道については、丸太材積に占める A 材、B 材、C 材の材積比率、さらに B 材、C 材を合わせた BC 材比率を算出し、これらを説明変数の候補とした。なお、本稿において A 材は「一般材」、B 材は「低質材」、C 材は「原料材」と整理することとする。

④伐採方法比率（皆伐、択伐、複層伐、間伐）

立木販売の対象には、皆伐や間伐の他、複層伐なども含まれ、この伐採方法の違いも、販売額に影響を与えている可能性がある。後に全物件の件数別比率などを示すが、多くの販売物件は単独の伐採方法からなることが多い。しかし、複数の伐採方法が混在していることから、各物件について伐採方法別の比率を変数化した。伐採方法は、「国有林野情報管理システム」においてはより細かく分類されているが、例えば間伐物件はそもそも件数が少なく、細分化することでさらに件数が減ることから、皆伐、択伐、複層伐、間伐の4つに分類した上で、それぞれの割合を算出した。なお、定性間伐と列状間伐では作業条件が異なることが予想されるが、生産変動経費や面積、材積等のデータが変数に含まれており、これらのデータ自体に対象物件の生産性等も反映されると考えられる。

⑤応札者数

推定に用いる販売物件データは販売済みであるから、販売の結果が得られるが、そこから応札者数を変数化した。応札者数が多い場合には、予定価格に対して高い販売額になることが予想される。

ただし、樹木採取権制度では、公募時に樹木料の算定方法を明らかにすることとしており、その時点で応札者相当の数値である申請者の数は得られない。また、仮に樹木採取権の申請者の数を式に適用することが可能としても、申請者の数によって樹木採取権毎に樹木料が変動することは、公平性に欠けるとも考えられる。そのため、応札者数を変数として用いないこともありえたが、分析を進める中で、やはり応札者数が販売額にはっきりした影響のあることが捉えられたため、説明変数候補として採用した。この場合、予測に当たっては、一律に過去のデータでの応札者数の平均を当てるという使用方法が考えられる。

⑥林地区分ダミー（直営林、分収造林、分収育林）、ブロックダミー、年度ダミー

販売物件の置かれた状況の違いを表す3つのダミー変数（群）である。まず、既述の通り、販売物件は直営林、分収造林、分収育林のいずれかであるため、これについてダミー変数を用意した。ダミー変数とは、材積や面積のように数値で表される量的データではない、質的データのうちカテゴリーを表す変数として表したものであり、具体的には、あるカテゴリーに当てはまる場合に1を、当てはまらない場合に0を取る変数である。

地域差を表すダミー変数として、全国を38に分ける販売ブロックのダミー変数を用意した。販売ブロックは、「販売ブロック運営要領の制定について（平成6年3月11日6林野業一第2号）」に元づき、各森林管理局長が定めた区域の単位であり、この単位で、木材需給事情、市況の動向等の実態の調査分析がされ、先の丸太価額を算定する際に使う素材の市況調査もこの単位で実施されている。そのため、一体的な素材の市場、流通圏の地域的広がりを表すのに適当だと考えた。

市況等、立木販売を取り巻く状況は時間とともに変化していることから、年度のダミー変数を採用した。なお、市況については、木材価格統計の月次・地域別の樹種別素材価格を接続して用いることも試みたが、有意な係数推定値は得られなかった。市況は、丸太価額、丸太単価の変化に反映されることで販売額への影響は捉えられると判断し、結果的にそれ以外の変数を入れることはしないこととした。時系列の変化に関しては、他にも季節ダミー変数やタイムトレンドも試したが、有効ではないと判断し、最終的には説明変数の候補から外した。

⑦森林調査簿等のデータ

上述したデータ以外に、販売された物件の内容についてのデータとして、森林調査簿のデータを林小班番号等で接続して変数化することも試みた。これによって、例えば、販売物件を構成する各林分の傾斜や標高、林道距離などに関するデータが得られる。しかしながら、林道距離の区分が森林管理局によって異なっていたり、森林調査簿の作成時点と実際の販売時期との違いから正しいデータが得られないものがあったりした。さらに、実際に回帰式の推定に使ってみてもあまり有意な影響が認められなかったことから、森林調査簿のデータの変数化は行わないこととした。

表3、4は被説明変数である販売額と、説明変数のうち⑥林名区分、ブロック、年度のダミーを除く①～⑤のグループの変数の記述統計を本州以南の局と北海道局でそれぞれまとめたものである。なお、単位の記載のないものは、全て0から1の間を取る比率で無名数である。

表3 使用した変数の記述統計（本州以南の局）

変数	単位	N	平均	中央値	最小値	最大値
販売額	円	3,411	6,718,196	4,500,000	3,000	63,700,000
丸太価額	円	3,411	21,673,455	16,254,561	67,281	205,715,188
生産変動経費	円	3,411	17,334,105	12,441,168	63,108	226,919,084
生産固定経費	円	3,411	2,093,343	1,746,999	1,155	13,503,783
丸太単価	円/m ³	3,411	9,524	9,675	2,574	18,639
生産変動単価	円/m ³	3,411	7,940	7,428	2,323	35,442
生産固定単価	円/m ³	3,411	1,356	1,107	4	70,088
面積	ha	3,411	6.10	4.67	0.01	60.33
立木材積	m ³	3,411	2,688	2,041	12	27,558
単位蓄積	m ³ /ha	3,411	462	458	8	1,327
単木材積	m ³ /本	3,411	0.54	0.47	0.03	3.56
本数密度	本/ha	3,411	973	933	19	4,170
林齢	年	3,411	55.69	56	25	119
スギ材積比率		3,411	0.68	0.84	0.00	1.00
ヒノキ材積比率		3,411	0.14	0.00	0.00	1.00
カラマツ材積比率		3,411	0.06	0.00	0.00	1.00

トドエゾ材積比率		3,411	0.00	0.00	0.00	0.08
その他針材積比率		3,411	0.06	0.00	0.00	1.00
広葉樹材積比率		3,411	0.06	0.02	0.00	1.00
平均胸高直径	cm	3,411	24.85	24.40	8.04	67.20
平均樹高	m	3,411	17.28	17.12	6.05	30.18
低質材率		3,411	0.17	0.10	0.00	1.00
一般材 (A 材) 比率		—	—	—	—	—
低質材 (B 材) 比率		—	—	—	—	—
原料材 (C 材) 比率		—	—	—	—	—
皆伐材積比率		3,411	0.92	1.00	0.00	1.00
択伐材積比率		3,411	0.00	0.00	0.00	0.04
複層伐材積比率		3,411	0.01	0.00	0.00	1.00
間伐材積比率		3,411	0.08	0.00	0.00	1.00
応札者数	人	3,411	2.27	2.00	1.00	11.00

表 4 使用した変数の記述統計 (北海道)

変数	単位	N	平均	中央値	最小値	最大値
販売額	円	970	3,038,160	1,787,613	2,000	33,080,000
丸太価額	円	970	11,984,596	10,013,678	60,244	77,495,527
生産変動経費	円	970	9,655,365	8,141,566	75,311	46,877,237
生産固定経費	円	970	2,458,045	1,947,591	67,673	19,848,583
丸太単価	円/m ³	970	9,277	9,318	6,680	13,124
生産変動単価	円/m ³	970	7,695	7,678	3,954	14,648
生産固定単価	円/m ³	970	2,231	1,940	256	20,971
面積	ha	970	22.46	17.02	0.2	195.36
立木材積	m ³	970	1,499	1,250	9	9,348
単位蓄積	m ³ /ha	970	100	65	3	634
単木材積	m ³ /本	970	0.43	0.34	0.07	2.66
本数密度	本/ha	970	280	203	14	1,379
林齢	年	970	51	52	24	145
スギ材積比率		970	0.01	0.00	0.00	1.00
ヒノキ材積比率		970	0.00	0.00	0.00	0.00
カラマツ材積比率		970	0.18	0.00	0.00	1.00
トドエゾ材積比率		970	0.67	0.84	0.00	1.00
その他針材積比率		970	0.01	0.00	0.00	1.00
広葉樹材積比率		970	0.13	0.08	0.00	0.99
平均胸高直径	cm	—	—	—	—	—
平均樹高	m	—	—	—	—	—
低質材率		970	0.00	0.00	0.00	0.00
一般材 (A 材) 比率		970	0.65	0.65	0.26	0.96
低質材 (B 材) 比率		970	0.20	0.21	0.02	0.44

原料材（C材）比率		970	0.15	0.15	0.00	0.36
皆伐材積比率		970	0.19	0.00	0.00	1.00
択伐材積比率		970	0.01	0.00	0.00	1.00
複層伐材積比率		970	0.06	0.00	0.00	1.00
間伐材積比率		970	0.74	1.00	0.00	1.00
応札者数	人	970	2.4	2.0	1.0	9.0

⑥のグループのダミー変数化したもののうち、林名区分と年度別の件数については、既に表1、2で紹介した。表5は販売ブロック別の販売物件数を見たものである。

表5 推定に使用した物件の販売ブロック別内訳

ブロック	件数	%	ブロック	件数	%	ブロック	件数	%
札幌	110	2.5	浜通り	116	2.7	岐阜南部	26	0.6
旭川	102	2.3	中通り	193	4.4	愛知	27	0.6
北見	410	9.4	南会津	68	1.6	紀伊	4	0.1
帯広	274	6.3	栃木	140	3.2	北陸京滋	1	0.0
函館	74	1.7	群馬	168	3.8	瀬戸内	51	1.2
青森西部	187	4.3	新潟	30	0.7	山陰	3	0.1
青森東部	181	4.1	茨城	143	3.3	四国	89	2.0
北岩手	231	5.3	東京	16	0.4	北九州	50	1.1
南岩手	213	4.9	静岡	41	0.9	熊本	91	2.1
宮城	113	2.6	富山岐北	19	0.4	大分	45	1.0
秋田県北	235	5.4	東北信	31	0.7	宮崎	345	7.9
秋田県南	159	3.6	中南信	6	0.1	鹿児島	165	3.8
山形	213	4.9	木曽	11	0.3	計	4,381	100.0

樹種、伐採方法は、北海道と本州以南ではかなり異なっていた。樹種について、表6は全販売物件の樹種比率の平均を見たものである。本州以南ではスギ、ついでヒノキが多く、北海道では、トドエゾ、ついでカラマツが多い。続いて、表7は、各樹種が立木材積で90%以上を占める物件の件数を示したものである。その他に分類されたものはどの樹種も90%未満の物件であるが、これが本州以南でも北海道でもほぼ半数を占めている。残りの半数はいずれかの樹種が90%以上を占めるもので、地域ごとの樹種の特徴は表6と同じである。ただし、北海道では、広葉樹の比率が比較的高いが、広葉樹ばかりで構成される物件は少なく、針葉樹と一緒に販売されることが多いものと見られる。

表6 平均樹種比率

	スギ	ヒノキ	カラマツ	トドエゾ	その他針	広葉樹
本州以南	0.685	0.139	0.059	0.000	0.056	0.062
北海道	0.013	0.000	0.181	0.667	0.012	0.127

表 7 各樹種が 90%以上を占める物件数

	90%以上						その他	計
	スギ	ヒノキ	カラマツ	トドエゾ	その他針	広葉樹		
本州以南	1,462	195	33	0	14	11	1,696	3,411
北海道	10	0	55	393	1	1	510	970

同様に、伐採方法について、表 8 で全物件の平均比率、表 9 で各伐採方法が 100%の物件数を見た。本州以南では皆伐が 90%を超えて圧倒的に多いが、北海道では間伐が 70%程度で多く、また択伐や複層伐の物件も少数ではあるが存在する。表 9 でその他が少ないように、販売物件のほとんどはいずれかの伐採方法単独のものであるが、特に北海道では 1 割強が複数の伐採方法が混じった販売の仕方となっている。このため、伐採方法については 0、1 のダミー変数を用いて分析することも考えられたが、立木販売の実績においては複数の伐採方法が混在した契約物件が（得に北海道局において多く）見受けられたため、これらのデータも含めて分析が可能であることから、既述の通り、伐採方法比率を変数化した。加えて、仮に 1 つの樹木料の算定単位（伐区）で複数の伐採方法が混在するようなことがあった場合に、伐採方法比率を変数化する方法であればどのような伐採方法の組合せであっても算定が可能であることから、今後の樹木料算定の運用面からも適当であると考えられた。

表 8 平均伐採方法比率

	皆伐	択伐	複層伐	間伐
本州以南	0.916	0.000	0.007	0.076
北海道	0.188	0.014	0.063	0.735

表 9 各伐採方法が 100%の物件数

	100%				その他	計
	皆伐	択伐	複層伐	間伐		
本州以南	3,111	0	24	257	19	3,411
北海道	166	10	26	648	120	970

最後に、被説明変数である販売額の分布について確認しておこう。図 1 は、全国の販売額のヒストグラムである。階級幅を 500 千円として描いてある。販売額は 1、2 千円から 60 百万円台まで幅広いが、安価な価格帯ほど件数が多く、右裾の長い分布となっている。全国の販売額の平均は図に書き込んだように 5,903 千円、中央値は 3,700 千円で、平均が中央値の約 1.6 倍であった。

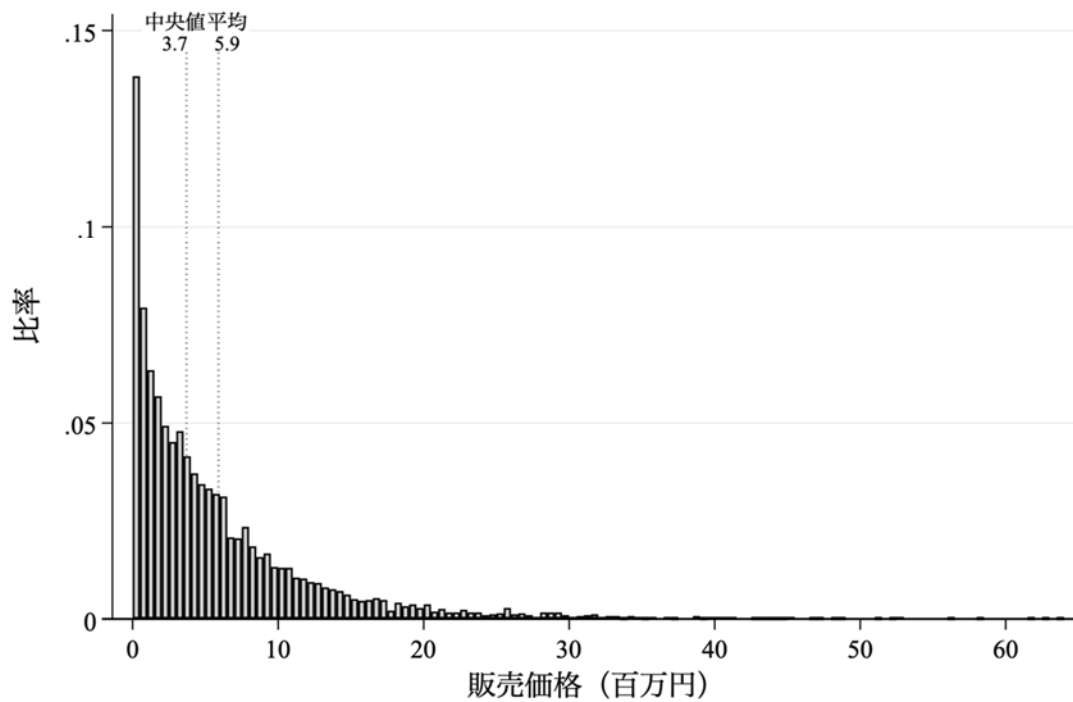


図1 販売額のヒストグラム

図2は販売額を立木材積で除して求めた販売単価（円/m³）のヒストグラムである。販売単価は5,000円以内にある。具体的には、販売単価の平均は2,206円、中央値2,100円、標本の90%は270円から4,618円の中にある。

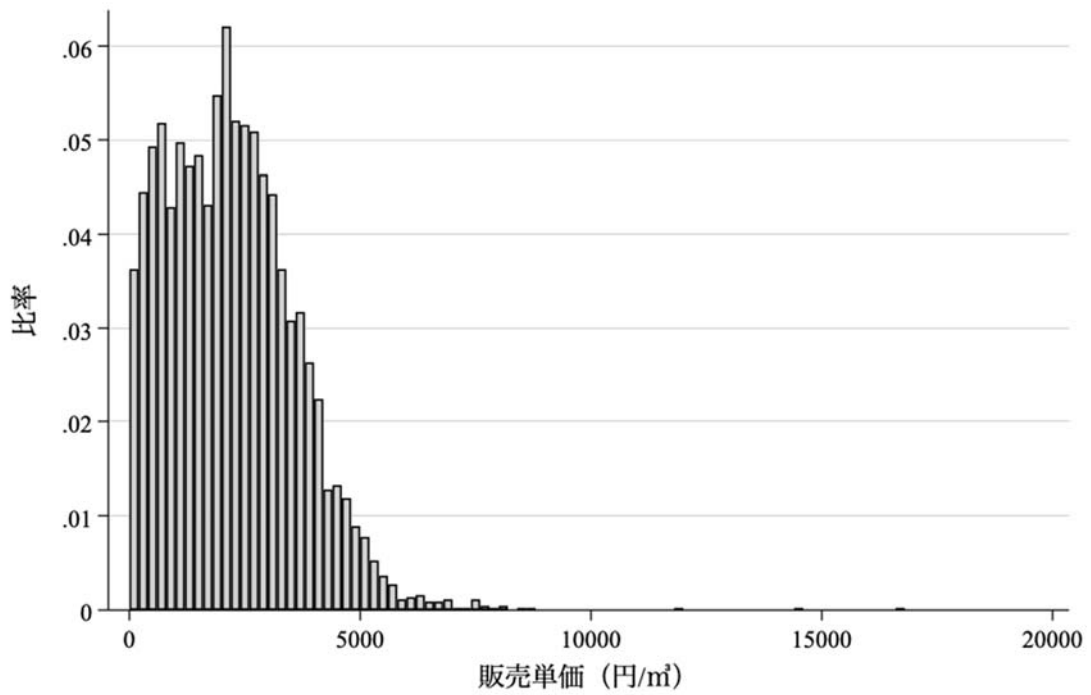


図2 販売単価のヒストグラム

図3は、販売額と予定価格との関係を散布図に描いたものである。販売額は必ず予定価格を上回るので、各物件を表す点は全て45度線より左上に位置する。点が重なって見にくいですが、多くの点は予定価格に近いところに分布するが、予定価格から大きく離れて高額で落札されるものが少数あることが見て取れる。

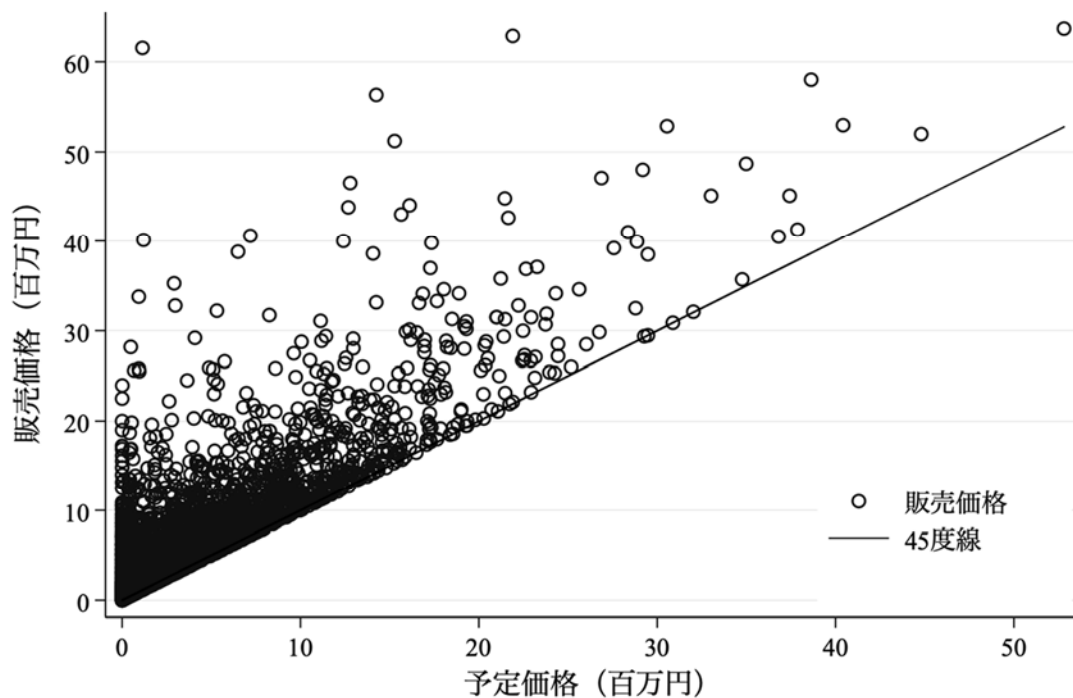


図3 予定価格と販売額

(4) 推定結果

表10に回帰式の推定結果をまとめた。被説明変数は販売額であるが、回帰式の当てはまりの良さを見ながら、最終的には対数変換した $\ln(\text{販売額})$ を用いた。回帰式自体は1次式でしか表されないため、対数変換等、変数をあらかじめ加工することで式の当てはまりが向上することがある。また、こうして対数変換を利用することで、説明変数がどんな値を取り、 $\ln(\text{販売額})$ がマイナスを取ろうとも、販売額の予測値は必ず正になることもメリットであった。他方、被説明変数を対数変換すると、対数変換後の変数の予測精度が高いモデルを選んでも必ずしも変換前の変数の予測精度が高いとは限らないとのデメリットもある。そこで、後に見るように変換前の販売額の予測精度を残差を用いて検討した上で、最終的に対数変換を用いて回帰を行うこととした。また、説明変数のうち、予定価格を通して販売額を規定する強い関係にある丸太価額、生産変動経費、生産固定経費についても当てはまりを見ながら、最終的に対数変換を用いることとした。

説明変数のうち、ブロックと年度、林名区分にダミー変数を使用した。具体的には、ブロックについては、北海道は函館ブロック、本州以南では鹿児島ブロックを完全な多重共線性を排除するために除外し、その他の全ブロックのダミー変数を用いた。年度については、2014年度を除く各年度のダミー変数を用いた。林名区分については、直営林を除く、分収造林と分収育林のダミー変数を用いた。ここで、完全な多重共線性とは、複数の説明変数 X_1, X_2, \dots の間で線形関係 $a_1X_1+a_2X_2+\dots=b$ が全ての標本について常に成り立っていることを意味する。例え

ば、林名区分に関して、直営林ダミー X 直、分収造林ダミー X 造、分収育林ダミー X 育については、 X 直 $+X$ 造 $+X$ 育 $=1$ が常に成り立つ。この場合、3つのダミー変数と定数項を同時に回帰式の説明変数として使うと、係数を推定できなくなる。

また、樹種比率、伐採方法比率、北海道局の丸太材積における一般材、低質材、原料材の材積比率についても、完全な多重共線性を防ぐため、少なくとも1つの比率を落として使用した。樹種比率の場合、北海道では、ヒノキを含む物件がなく、その他針葉樹を含む物件もわずかであったため、この2つの比率を除いた上に、さらにスギ比率を落として、他樹種の比率を推定に用いた。本州以南の局では、トドエゾを含む物件がなかったため、その比率を除いた上に、その他針葉樹比率を落として、他樹種の比率を推定に用いた。伐採方法比率は、北海道局、本州以南の局とも皆伐材積比率を落とし、択伐、複層伐、間伐の材積比率を用いた。北海道の丸太材積における一般材、低質材、原料材の材積比率については一般材材積比率を落とし、低質材材積比率、原料材材積比率を用いた。また、既述の通り、低質材、原料材合計の材積比率の使用も試みたが、Stepwiseで除去されたため、最終的には、低質材材積比率、原料材材積比率を用いた。

表10では、推定はOLS回帰については、StepwiseをP値基準で行ったものとAIC基準で行ったものの両方を掲載した。中央値回帰はP値基準のみである。それぞれのケースについて、Stepwise変数減少法により残った変数についてだけ係数とその有意性を示した。表に登場しない変数はいずれの推定からもStepwiseで除外されたものである。

OLS回帰について結果を見ると、まず、P値基準とAIC基準では、極めて似通った推定結果となった。Stepwiseで残った変数もかなり共通しており、その係数推定値も似た値を取るものが多い。決定係数 R^2 もほぼ同じである。従って、P値基準による変数選択もAIC基準による変数選択もほぼ同じ推定結果にたどり着いたと言える。

R^2 は本州以南で0.87、北海道で0.85程度とかなり説明力の高い結果となっている。係数の符号に関しては、丸太価額、生産変動経費、生産固定経費の対数変換の係数が、丸太価額について正（プラス）、生産変動経費、生産固定経費について負（マイナス）となっていることは、妥当な結果と言える。一方で、例えば、本州以南の対数変換された丸太価額がプラスであるのに対し、丸太単価の符号はマイナスとなっている。このように、一見矛盾するような係数と符号の組合せも見受けられるが、今回の分析の目的は販売額の予測であり、Stepwise法により最も当てはまりの良い式を作成することであることから、結果的に、より高い精度で販売額を予測できるのであればそれでよい。また、回帰式は一次式で表されることから、相反する符号の変数があることで相殺による調整がされているともとらえることができ、符号の逆転現象によって算定式に瑕疵があるということにはならない。

中央値回帰の結果については、選ばれた変数はOLS回帰と類似しており、その係数推定値はOLS回帰といくらか異なるが、符号など似た傾向の結果となった。全体の当てはまりを見る疑似 R^2 は本州以南で0.69、北海道で0.64程度であった。この疑似 R^2 とOLS回帰の R^2 を直接比

べることはできず、OLS 回帰と中央値回帰の予測精度の違いなどについては、項を改めて残差の分析を行う。

なお、OLS 回帰については、VIF (Variance Inflation Factor) を算出し、説明変数間の多重共線性について検討した。説明変数間に多重共線性がある場合、係数の推定値が真の値とはかけ離れた値を取る可能性が高まるとの問題が生じる。その結果、本州以南でも、北海道でも、ln(A 価額)と ln(B 経費)の VIF が 10 を超えており、両変数間には強い相関が見られたが、その他の変数では VIF が 10 を超えるものはなかった。丸太価額と生産変動経費は販売額を規定する要因として最重要のもので、式から除外することは考えにくい。また、丸太価額と生産変動経費が強い相関を持つことは計算方法上想定されることで、今後予測に用いる物件においても同じ関係が維持されることは確実であるから、予測が主目的の今回の回帰式推定では、多重共線性によって係数推定の不安定性が生じていたとしても、両変数を残しておいて問題はないと判断した。

表 10 回帰式推定結果

	被説明変数：ln(販売額)					
	OLS 回帰				中央値回帰	
	P 値基準		AIC 基準		P 値基準	
	本州以南	北海道	本州以南	北海道	本州以南	北海道
ln(丸太価額)	1.633190 ***	3.918766 ***	1.567681 ***	3.300501 ***	1.714857 ***	3.066645 ***
ln(生産変動経費)	-0.581000 ***	-2.473770 ***	-0.580683 ***	-1.850225 ***	-0.633930 ***	-1.605275 ***
ln(生産固定経費)	-0.051794 *	-0.218092 ***	-0.042768 *	-0.213330 ***	-0.096656 ***	-0.165563 ***
丸太単価	-0.000070 ***		-0.000063 ***		-0.000052 ***	
生産変動単価		0.000090				-0.000049
生産固定単価			-0.000010			
面積		-0.028168	0.014798 *	-0.031632 *		-0.028931 ***
面積 10ha 超	-0.008981 *	0.024433	-0.022564 **	0.027555	-0.005839	0.028748 ***
立木材積						-0.000100 ***
林齢	0.003059	0.004300 *	0.003245 *	0.004235 *	0.003032 ***	0.001836
単木材積	0.271993 *	-0.297546 *	0.217537 ***	-0.204749	0.144078 ***	-0.206107 ***
本数密度	0.000055	-0.000486 **	0.000077 *	-0.000491 **		-0.000149
平均樹高	-0.011709				-0.007336 *	
分収造林					-0.025729	
分収育林	0.164241 ***		0.166949 ***		0.083307 ***	
択伐材積比率	30.032824 ***		30.664507 *			
複層伐材積比率	-0.342516 *		-0.373621 ***		-0.405127 ***	0.214674 **
間伐材積比率	-0.731230 ***		-0.789403 ***		-0.662570 ***	
スギ材積比率	0.539591 ***		0.547731 ***		0.318131 ***	
ヒノキ材積比率	0.829865 ***		0.807180 ***		0.509964 ***	
カラマツ材積比率	1.237049 ***	0.226511	1.213309 ***	0.320415	0.942997 ***	0.480126 ***
トドエゾ材積比率		-0.458409 *		-0.311464 *		
広葉樹材積比率		-1.151736 ***		-0.813215 ***	-0.371195 *	-0.708118 ***
低質材材積率	-0.600258 ***		-0.578809 ***		-0.201983 **	
低質材 (B 材) 比率						-0.884630
原料材 (C 材) 比率		2.425605 **		1.979459 **		2.187778 ***

応札者数	0.128790 ***	0.169090 ***	0.128723 ***	0.169344 ***	0.094070 ***	0.126689 ***
2015年度	-0.103853 **		-0.103676 **		-0.067847 ***	
2016年度	-0.209628 ***	0.198582 ***	-0.205627 ***	0.194819 ***	-0.119980 ***	0.232221 ***
2017年度	-0.210847 ***	-0.370918 ***	-0.207653 ***	-0.384489 ***	-0.137964 ***	-0.433018 ***
2018年度	-0.196196 ***	-0.140605	-0.190959 ***	-0.147572	-0.128165 ***	-0.249471 ***
札幌						-0.096867
北見						0.151459 **
帯広		-0.066585				
青森西部	-0.244837 ***		-0.229252 ***		-0.179832 ***	
青森東部	-0.458646 ***		-0.447153 ***			
北岩手					0.078038	
南岩手	-0.359803 ***		-0.352754 ***		-0.288030 ***	
宮城					0.146716 ***	
秋田県北	-0.226577 ***		-0.231427 ***		-0.111968 ***	
秋田県南	-0.186777 ***		-0.169493 ***		-0.070125 *	
山形	-0.228360 ***		-0.228576 ***		-0.116032 ***	
浜通り	0.473852 ***		0.486597 ***		0.599411 ***	
中通り	0.241347 ***		0.245573 ***		0.422327 ***	
栃木	-0.088761		-0.084272		0.149070 **	
群馬	-0.281692 ***		-0.278580 ***		-0.084457 *	
新潟	-0.178720		-0.176785		-0.179696 ***	
茨城					0.148245 ***	
東京	-0.300109		-0.293411 *		-0.142828 ***	
静岡	-0.178652		-0.163284			
富山岐北	-0.335495 ***		-0.327939 *			
東北信	-0.396059 ***		-0.394919 ***		-0.250632 ***	
中南信	-0.318534 *					
木曾	-0.396682		-0.374393 *			
岐阜南部	-0.084940					
愛知					0.198776 ***	
紀伊					0.129838	
北陸京滋	-0.135944					
瀬戸内					0.056749	
山陰	-0.169448 *					
四国	-0.229625 ***		-0.229360 ***		-0.138320 ***	
北九州	-0.076667				0.045514	
熊本					0.057647 *	
大分					0.093077 ***	
宮崎					0.074492 ***	
定数項	-1.640227 ***	-7.031084 ***	-1.050538 **	-6.455873 ***	-1.451092 ***	-6.731350 ***
N	3411	970	3411	970	3411	970
R ²	0.8715	0.8508	0.8714	0.8503		
疑似 R ²					0.6957	0.6364

注：***は P 値 ≤ 0.001、**は ≤ 0.01、*は ≤ 0.05。P 値基準の場合、OLS 回帰、中央値回帰とも分散共分散行列の推定にはロバスト推定を行った。

(5) 予測残差の分析

OLS 回帰と中央値回帰の結果について、予測式としての予測精度という面で比較検討するため、予測残差の分析を行った。ここで、予測残差とは、回帰式の推定結果を使って予測した販売額予測値から、販売額の実現値を引いて得た残差のことである。なお、回帰式は $\ln(\text{販売額})$ を被説明変数とする式であるので、回帰式を使って予測した $\ln(\text{販売額})$ を指数変換して、販売額の予測値とし、実現値との差を取っている。

表 11 は表 10 の回帰式推定結果を用い、残差を得て、その絶対値、また平方を取って全物件での平均を見たものである。表頭の全体は推定に用いた標本全ての予測残差についての平均であり、直営林のみは推定に用いた標本のうち直営林のみの平均である。樹木料採取権制度では分収林は対象外であるため、直営林のみの集計も見ることとした。

1 物件当たりの平均残差が小さいほど販売額を精度高く予測できているといえるが、結果を見ると、絶対残差でも平方残差でも、また、北海道局でも本州以南の局でも、推定標本全体でも直営林に限定しても、中央値回帰の方が平均残差が小さいとの結果となった。推定標本全体について見ると、OLS 回帰では、平均残差は P 値基準でも AIC 基準でもほぼ変わらず、全国の絶対残差で 156~157 万円、平方残差の平方根で 297~298 万円であるのに対し、中央値回帰では絶対残差の平均が 139 万円、平方残差の平方根で 261 万円と、中央値回帰の方がいずれも 13%程度ずつ残差が小さかった。直営林のみで見ても同様に、中央値回帰の方が 10%強平均残差が小さい結果となった。

OLS 回帰は平方残差和 (= 残差平方和) を最小にするよう係数を選ぶ方法、中央値回帰は絶対残差和を最小にするように係数を選ぶ方法であることからすれば、平均絶対残差和は中央値回帰の方が小さく、平均平方残差の平方根は OLS 回帰の方が小さくなること予想されるが、Stepwise による変数選択をしており、最終的に採用した変数の組み合わせが OLS 回帰と中央値回帰で異なる上、被説明変数そのものではなく、指数変換した販売額で残差を見ているため、そうはならず、いずれも中央値回帰の方が小さいとの結果となった。

また、OLS 回帰の P 値基準と AIC 基準ではほぼ結果は変わらず、全体としては、わずかに P 値基準の残差の方が小さかったので、これ以降は、OLS 回帰、中央値回帰統一して P 値基準での変数選択結果についての検討を行う。

表 11 平均残差による比較

	回帰式	Stepwise 変数選 択基準	全体		直営林のみ	
			平均絶対残 差	平均平方残 差の平方根	平均絶対残 差	平均平方残 差の平方根
本州以南の局 (N=3,411)	OLS 回帰	P 値	1,711,668	3,163,428	1,276,322	2,209,667
		AIC	1,712,168	3,173,739	1,281,506	2,243,821
	中央値回帰	P 値	1,543,879	2,831,991	1,184,637	2,098,854
北海道局 (N=970)	OLS 回帰	P 値	1,028,069	2,329,344	1,033,370	2,128,525
		AIC	1,000,904	2,179,483	1,009,596	2,008,105
	中央値回帰	P 値	852,760	1,513,185	858,014	1,363,825
全国 (N=4,381)	OLS 回帰	P 値	1,560,312	2,998,816	1,158,386	2,170,657
		AIC	1,554,687	2,982,309	1,149,513	2,132,654
	中央値回帰	P 値	1,390,858	2,598,344	1,026,085	1,780,361

残差については、別の方法で残差を得て比べることも行った。クロスバリデーションの方法を用いるもので、OLS 回帰、中央値回帰とも、表 10 の P 値基準の Stepwise で変数選択済みのモデルについて、一つは、(I) 標本から 1 物件ずつを除いて係数を推定し直し、推定された式を用いて推定から除いた物件の予測値を得て残差を得る方法、もう一つは、(II) 標本から 1 年度分ずつを除いて推定し直し、除いた年度の物件の予測値を得て残差を得る方法である。

表 12 はクロスバリデーション (I) の方法で得た全国の平均残差であるが、結果的に、表 11 の全国の部分とあまり変わらない結果となった。今回標本数が北海道でも 1 千を超え十分に大きいため、1 物件ずつ除いて推定を行うクロスバリデーションの方法でも、推定結果は全標本を用いた場合とほぼ変わらず、得られる残差にも大きな差は見られなかった。OLS 回帰と中央値回帰それぞれの本州以南と北海道ごとの 4 本の式について、それぞれ得た通常残差とクロスバリデーション (I) の残差の相関係数は全て 0.998 以上であった。その結果、表 12 に見るように、クロスバリデーションによる平均残差は表 11 で示した通常残差と比べてやや大きくなるものの、中央値回帰の方が 10% 強平均残差が小さいとの結果は全体でも直営林限定でも変わらなかった。

表 12 クロスバリデーション (I) による平均残差による比較

	全体 (N=4,381)		直営林のみ (N=1,613)	
	平均絶対残差	平均平方残差の平方根	平均絶対残差	平均平方残差の平方根
OLS 回帰	1,583,512	3,044,096	1,181,757	2,219,947
中央値回帰	1,432,166	2,664,192	1,057,054	1,829,914

注：本州以南と北海道を合わせた 4,381 件の平均。P 値基準。

次に表 13 はクロスバリデーション (II) による平均残差の比較である。年度単位で標本を抜くため、残差は大きくなる。しかし、直営林のみの絶対残差では OLS 回帰と中央値回帰の差が

5%ほどに縮まるものの、OLS 回帰より中央値回帰の方が残差が小さいとの傾向はこれまでと同様であった。

表 13 クロスバリデーション (II) による平均残差による比較

	全体 (N=4,381)		直営林のみ (N=1,613)	
	平均絶対残差	平均平方残差の平方根	平均絶対残差	平均平方残差の平方根
OLS 回帰	1,884,136	3,655,638	1,414,077	3,053,638
中央値回帰	1,661,854	3,058,387	1,352,072	2,504,095

注：本州以南と北海道を合わせた全国の平均。P 値基準。

より詳しく残差の分布を見ることで OLS 回帰と中央値回帰の結果を比較するために、表 13 は、通常の残差を使い、本州以南と北海道それぞれについて、OLS 回帰と中央値回帰から得られた残差の度数分布を見たものである。なお、残差は予測値から実現値を引いて求めているので、残差が正であることは予測値が実現値を上回っていることを意味する。本州以南、北海道、全国計、いずれの比較においても、中央値回帰の方が 0 円の周辺への集中度がわずかずつではあるが高い。全国計では、残差が±100 万円（枠囲み部分）に収まる割合が OLS 回帰では 59.2%に対し、中央値回帰では 60.6%、±200 万円（二重線枠囲み部分）では OLS 回帰が 77.7%に対し中央値回帰が 79.6%などとなった。残差の分布形状には OLS 回帰と中央値回帰で明確な差は認められず、いずれの残差も中央が高くほぼ左右対象に分布していた。

表 13 残差の度数分布

	OLS 回帰						中央値回帰					
	本州以南		北海道		計		本州以南		北海道		計	
	度数	%	度数	%	度数	%	度数	%	度数	%	度数	%
~-8 百万円	65	1.9	1	0.1	66	1.5	54	1.6	2	0.2	56	1.3
~-6 百万円	39	1.1	3	0.3	42	1.0	28	0.8	3	0.3	31	0.7
~-4 百万円	105	3.1	7	0.7	112	2.6	77	2.3	5	0.5	82	1.9
~-2 百万円	301	8.8	55	5.7	356	8.1	273	8.0	33	3.4	306	7.0
~-1 百万円	373	10.9	106	10.9	479	10.9	324	9.5	89	9.2	413	9.4
~0 百万円	976	28.6	379	39.1	1,355	30.9	949	27.8	355	36.6	1,304	29.8
~1 百万円	926	27.1	314	32.4	1,240	28.3	997	29.2	354	36.5	1,351	30.8
~2 百万円	276	8.1	53	5.5	329	7.5	344	10.1	75	7.7	419	9.6
~4 百万円	192	5.6	25	2.6	217	5.0	228	6.7	36	3.7	264	6.0
~6 百万円	69	2.0	12	1.2	81	1.8	65	1.9	11	1.1	76	1.7
~8 百万円	46	1.3	4	0.4	50	1.1	30	0.9	5	0.5	35	0.8
8 百万円~	43	1.3	11	1.1	54	1.2	42	1.2	2	0.2	44	1.0
計	3,411	100.0	970	100.0	4,381	100.0	3,411	100.0	970	100.0	4,381	100.0

OLS 回帰と中央値回帰の予測精度について、残差を用いた検討を行ってきたが、次に、そもそも、OLS 回帰の予測値と中央値回帰の予測値がどう違うについて検討する。まず、図 4 は横軸に販売額の実現値を取り、縦軸に OLS 回帰と中央値回帰の予測値をプロットしたものである。予測値が 45 度線に近いほど、精度の高い予測ができていていることを示す。販売額が小さい部分は多くの点が集中して違いがわかりづらいが、販売額が 10 百万円を超えるあたりから、OLS 回帰の予測値が 45 度線の上方でより高いところに位置する場合が多くなっていることが目立つようになる。つまり、金額の大きな物件で OLS 回帰の過大予測が目立つのであり、このことが OLS 回帰が中央値回帰に対して、全体的な予測精度でやや劣る原因となっていることが考えられる。

実際に、図 5 で横軸に OLS 回帰の予測値、縦軸に中央値回帰の予測値をプロットしてみると、金額が大きい領域では、点が 45 度線より下にあり、中央値回帰の方が小さな予測をする傾向にあることが分かる。一方、少額な物件では、よく見ると、45 度線より上に点がやや分厚く集中している部分があり、中央値回帰の方が大きな予測をする傾向が見られた。

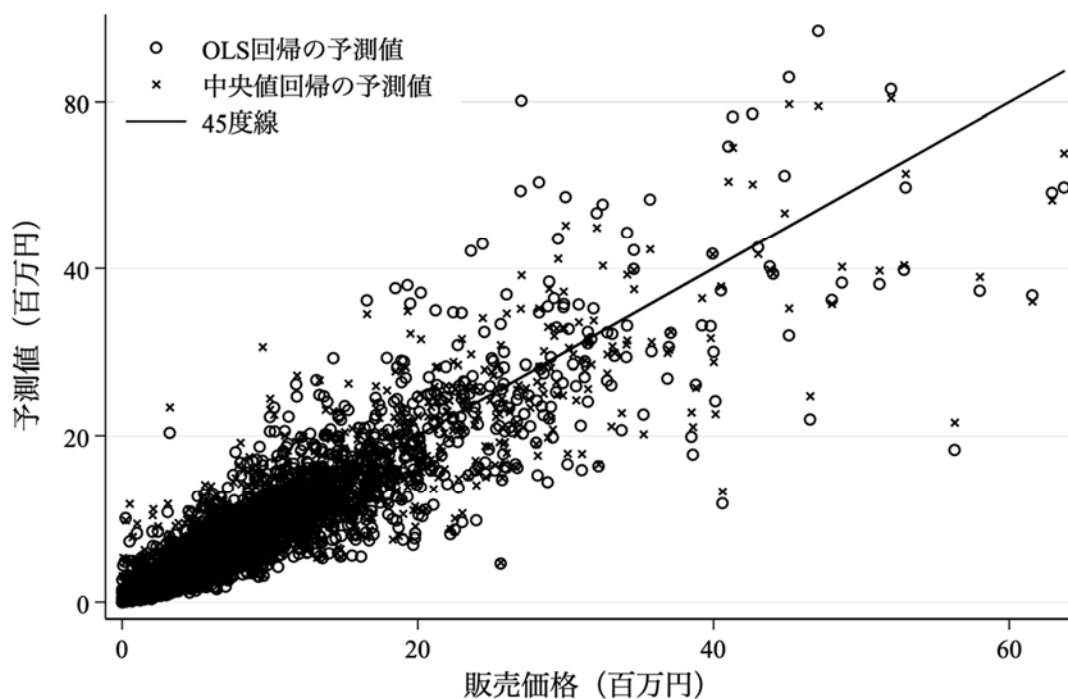


図 4 販売額と OLS 回帰と中央値回帰の予測値

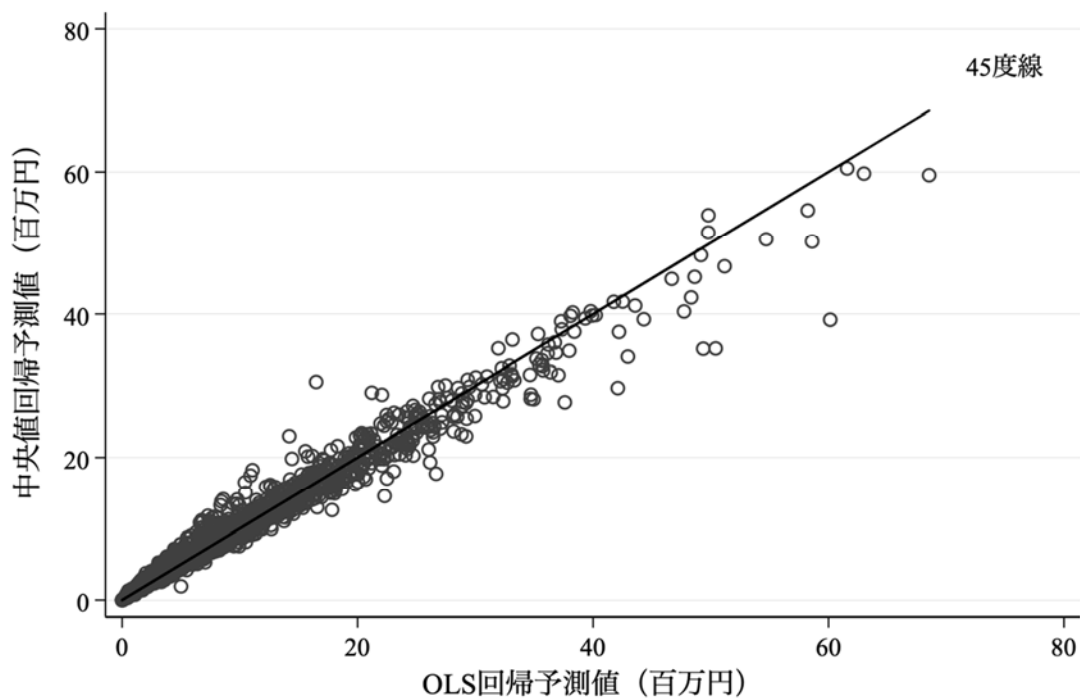


図5 OLS回帰と中央値回帰の予測値のプロット

このことと残差の関係を見るために、表14では販売額階層を分け、階層ごとにOLS回帰からの予測値と中央値回帰からの予測値との差、また、両回帰からの平均的な残差の大きさを見たものである。まず、「中央値回帰予測値-OLS回帰予測値」の平均の列では、販売額の各階層における2つの予測値の差の平均を見ているが、図にも表れていたように、販売額の小さな階層では、中央値回帰の方が100~300千円程度大きな予測をしているが、販売額の大きな階層では、OLS回帰の方が500~800千円程度大きな予測をしていることが分かる。そして、販売額階層ごとの平均絶対残差を見ると、販売額が小さな階層では、絶対残差の大きさはOLS回帰と中央値回帰の間で大差ないが、1百万円以下の階層でOLS回帰の方が残差が70千円程度小さいが、反対に販売額5百万円程度からは、明らかに中央値回帰の方が残差が小さい。絶対残差を販売額で割り、率にして平均を取った平均絶対残差率で見ても、この傾向は明らかである。このように、販売額が5百万円程度以上の金額が大きな物件における差が、計の行に再掲した全体の平均絶対残差の差を生んだと考えられる。従って、中央値回帰の方が全体として残差が小さく予測精度が高いのは、より大きな金額の物件についてOLS回帰が過大な予測をするのに対し、中央値回帰がより精度高く販売額を予測し得たからであると言える。その一方で、中央値回帰は全体の予測精度はOLS回帰と比べ相対的に高いが、1百万円以下の物件ではやや過大推定になっておりOLS回帰と比べても精度が劣ることは留意すべき点である。

表 14 販売額階層ごとの予測値、残差の比較

販売額階層	N	「中央値回帰予測値-OLS 回帰予測値」の平均	平均絶対残差		平均絶対残差率 ^(注)	
			OLS 回帰	中央値回帰	OLS 回帰	中央値回帰
～0.5 百万円	621	83,702	253,457	323,788	3.303	4.721
～1 百万円	351	136,963	420,431	495,422	0.561	0.664
～3 百万円	947	209,927	672,754	694,240	0.359	0.379
～5 百万円	709	256,061	1,091,420	1,022,265	0.277	0.26
～10 百万円	965	271,860	1,757,217	1,608,522	0.246	0.226
～20 百万円	584	12,735	3,275,961	2,690,298	0.237	0.194
～30 百万円	138	-782,478	6,122,553	4,661,851	0.25	0.192
30 百万円～	66	-530,226	10,091,763	8,628,153	0.248	0.211
計	4,381	138,600	1,560,312	1,390,858	0.733	0.931

注：残差率 = (予測値-販売額)/販売額。平均絶対残差率は、残差率の絶対値の各階層における平均。

方法のところで述べたように、OLS 回帰は被説明変数の分布の平均を、中央値回帰は中央値を予測しようとするのであった。そして、被説明変数である販売額は右裾の長い分布をしているから、この2つの回帰手法の特徴を考えれば、OLS 回帰の予測値は中央値回帰の予測値より大きいことが期待される。しかし、今回の結果では、階層計の「中央値回帰予測値-OLS 回帰予測値」の平均が正であるように、全体としては中央値回帰の予測値の方が大きくなっている。確かに、販売額が大きな部分では、中央値回帰の方がより小さい予測をし、予測精度が高いが、全体にわたって中央値回帰の方が OLS 回帰と比べて小さな予測をしているわけではない。従って、両回帰分析による予測結果の違いが、一方が平均を一方が中央値を予測しているから生じたことであると単純に理解することはできなかった。

では、なぜこのような予測結果の違いが生まれるのかが問題となるが、これについて、確たる説明を付けることはできなかった。中央値回帰の特徴として、中央値を問題にし、計算方法として絶対残差の和を最小にしようとすることから、OLS 回帰と比べ外れ値に強いということが言われる。確たる証拠は示し難いものの、このことが中央値回帰がやや良好な予測結果をもたらした理由である可能性はある。

5 検証

(1) 検証に使用したデータ

検証には分析に使用していない今年度の立木販売実績のデータを用いることとし、作業時点において抽出が可能であった2019年4月1日から11月15日現在までのデータを用いた。

なお、分析に使用したデータと同様に、検証に使用するデータも同様の条件（一般競争入札に限定し、混合契約や支障木等、樹木採取区の対象には含まれないことが判明したデータ等は除外）とした（表15）。

また、樹木採取権では分収育林及び分収造林は対象外であるため、これら分収林を除く計292件を検証対象データとした。

表15 検証に用いたデータ件数

局	直営林	分収育林	分収造林	計
北海道局	160	31	11	202
北海道以外の局	141	82	284	507
東北局	46	37	151	234
関東局	45	9	47	101
中部局	2	11	2	15
近中局	9	4	0	13
四国局	10	4	1	15
九州局	29	17	83	129
計	292	113	295	700

(2) 検証結果

分析結果では、OLS回帰（P値及びAIC）及び中央値回帰による方法が示され、中央値回帰による方法が最も当てはまりが良い結果となった。本節における検証では、OLS（P値）及中央値回帰の2パターンについて、算出額と実際の販売額の残差や残差の分布を比較することで、最終的にいずれの式を採用するかを決定することにした。

なお、樹木採取権制度では分収林は対象外であることから、検証対象から分収林を除くこととするが、式には分収林もダミー変数として含まれており、式の当てはまりを検証するにあたって、分収林を含めた場合の残差も示すこととする。

検証に用いたデータの残差の平均絶対残差を表16に、平均平方残差の平方根を表17に示す（ここでは参考までOLS（AIC）の結果も記している）。

今回の分析において応札者数に変数に含まれるが、実際の樹木料算定にあたっては応札者数を設定することはできない。そのため、検証に当たっては実際の応札者数を適用した場合と、

実際の樹木料評定額の算定を想定し、応札者数の平均値を一律で適用した場合について残差を算出することとした。

また、分析によって得られた2014年度から2018年度の各年度の係数については、検証及び実際の樹木料算定においては設定できないため、各年度の係数の平均値を用いることとした。

表 16 検証に用いたデータの平均絶対残差

	林名区分	件数	実際の応札者数を適用			応札者数は平均値を適用		
			ols_p	ols_aic	med_p	ols_p	ols_aic	med_p
北海道局	分収林含む	201	1,682,404	1,575,548	1,454,773	1,398,437	1,303,626	1,301,078
	分収林除く	159	1,535,701	1,453,131	1,387,672	1,383,456	1,305,561	1,310,741
本州以南の局	分収林含む	499	2,194,992	2,206,704	2,070,835	2,586,511	2,578,126	2,461,018
	分収林除く	133	1,533,068	1,561,517	1,312,120	1,707,319	1,717,900	1,520,031

表 17 検証に用いたデータの平均平方残差の平方根

	林名区分	件数	実際の応札者数を適用			応札者数は平均値を適用		
			ols_p	ols_aic	med_p	ols_p	ols_aic	med_p
北海道局	分収林含む	201	4,450,146	4,003,749	2,960,776	2,251,508	2,003,023	1,863,394
	分収林除く	159	2,862,233	2,653,622	2,195,486	1,944,627	1,796,452	1,737,672
本州以南の局	分収林含む	499	4,236,667	4,284,380	3,954,187	4,969,441	4,998,893	4,672,277
	分収林除く	133	3,034,920	3,063,753	2,634,282	3,358,824	3,387,322	2,980,191

※平均平方残差は各算出額と販売額の差（残差）の二乗の平均

※平均絶対残差は各算出額と販売額の差（残差）の絶対値の平均

※ols_p：OLS 回帰による P 値分析

※ols_aic：OLS 回帰による AIC 分析

※med_p：中央値回帰による P 値分析

実際の応札者数を適用した場合の残差を見ると、いずれの結果を見ても中央値回帰による残差が最も小さかった。

また、応札者数の平均値を適用した場合は、北海道局の分収林を除いた場合における平均絶対残差において、OLS（AIC）回帰による結果が最も小さかったが、それ以外は中央値回帰の残差が最も小さい結果となった。実際の樹木料算定では応札者数の平均値を適用するが、検証に使用するデータは応札者数が判明しており、検証の目的が算定式のあてはまりの良さを見ることであることから、実際の応札者数を適用した場合の残差を重視することとする。

以上を踏まえると、残差の比較等、算定式の検証に当たっては総合的に中央値回帰による回帰式の精度が最も良いことが裏付けられた。

次に、分取林を除く 292 件について、残差の分布を見てみる。

図 6～図 9 は OLS 回帰及び中央値回帰のそれぞれについて、実際の応札者数を適用した場合と、応札者数の平均値を適用した場合の残差の分布をヒストグラム化したものである。残差は算出額から実際の販売額を差し引いた額としているため、算出額が販売額よりも大きく算定される場合は正（プラス）、算出額が販売額よりも小さい場合は負（マイナス）となる。

表 16 及び表 17 から、平均残差は中央値回帰による方法が最も小さいことが明らかになったが、残差の分布をヒストグラム化した図 6～9 でも、中央値回帰が若干ではあるが、やや中央よりであることが見て取れる。また、物件によって、OLS 回帰及び中央値回帰のいずれにおいても極端に残差が大きい物件が見受けられる。例えば、算出額の倍の額で販売されている物件が存在するが、実際にこの物件について詳細を確認してみると、複数の応札があり、落札者とそれ以外の者の応札額の間でも同程度の差があることが判明した。これは極端な事例ではあるが、実際の販売額は、競争の結果、（特に多数の応札者による入札においては）競争原理が働き、実際の販売額が比較的高くなっていると推測される。応札者数は、分析において有意であることが明らかになっていることから、これとは逆に、応札者が 1 者の場合は、販売額が比較的安価になっている可能性も考えられる。

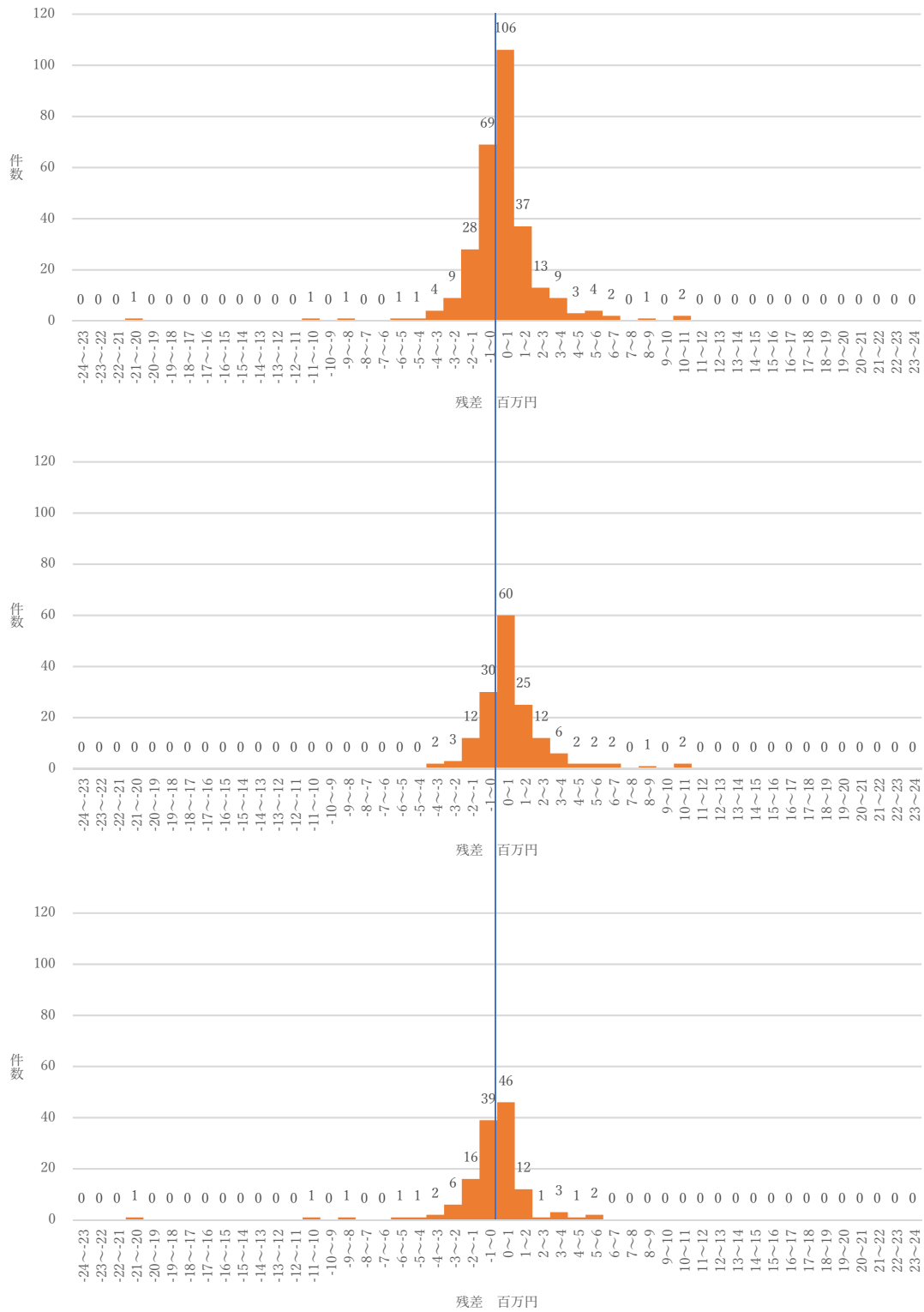


図6 中央値回帰の残差の分布（実際の応募者数を適用）
 （上：全局、中：北海道局、下：本州以南の局）

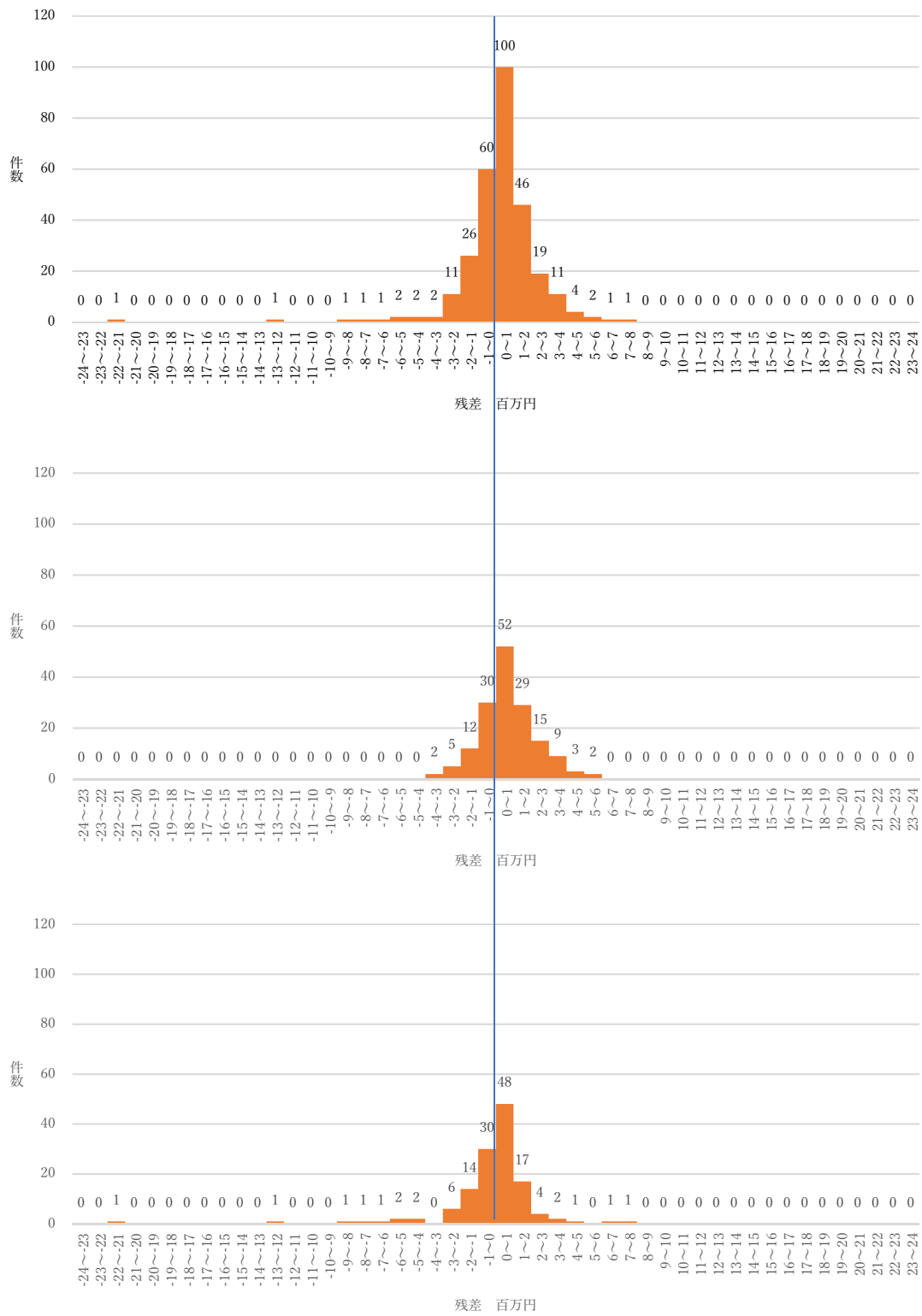


図7 中央値回帰の残差の分布（応札者数の平均値を適用）
 （上：全局、中：北海道局、下：本州以南の局）

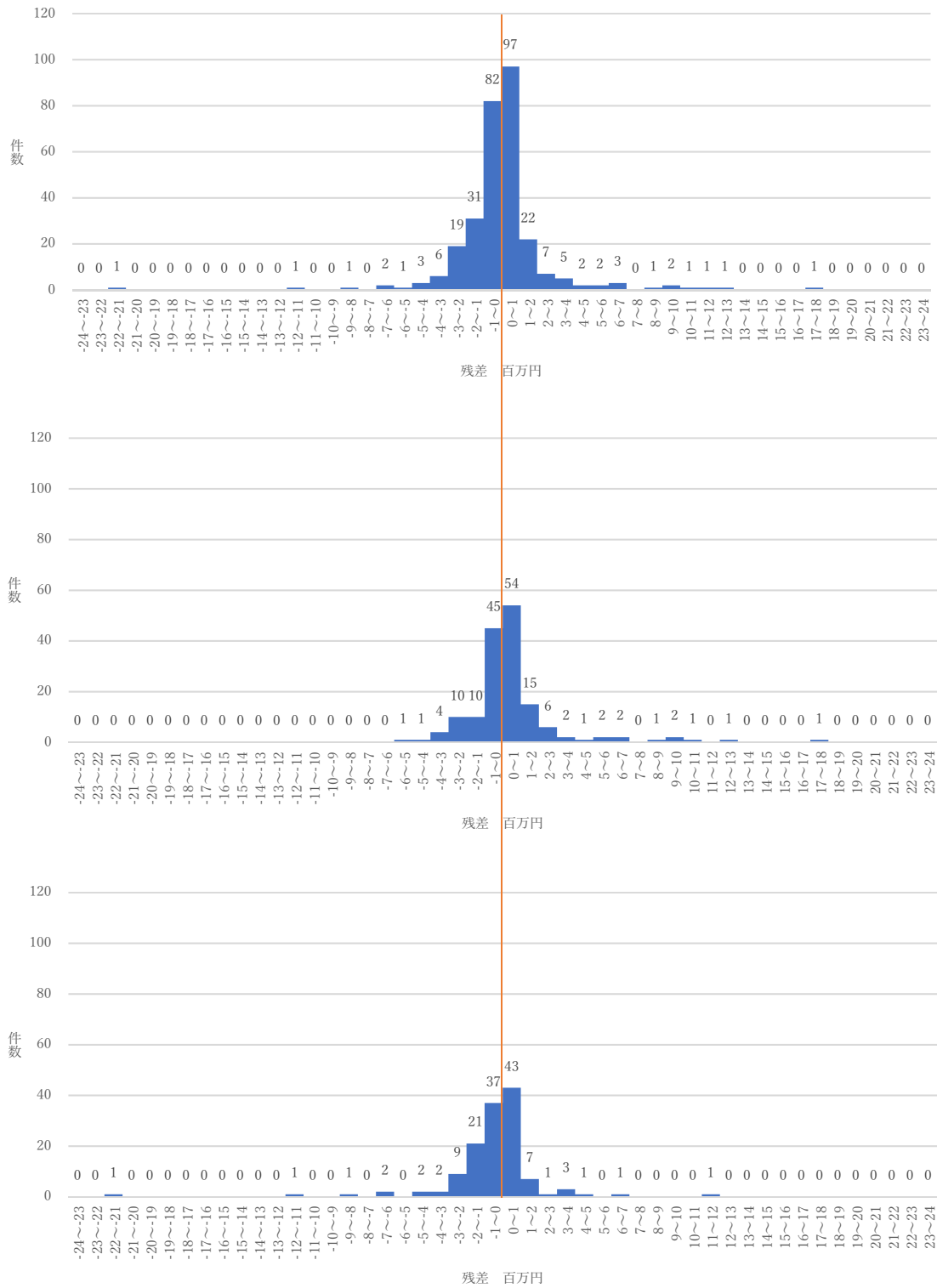


図8 OLS 回帰の残差の分布 (実際の応募者数を適用)
 (上: 全局、中: 北海道局、下: 本州以南の局)

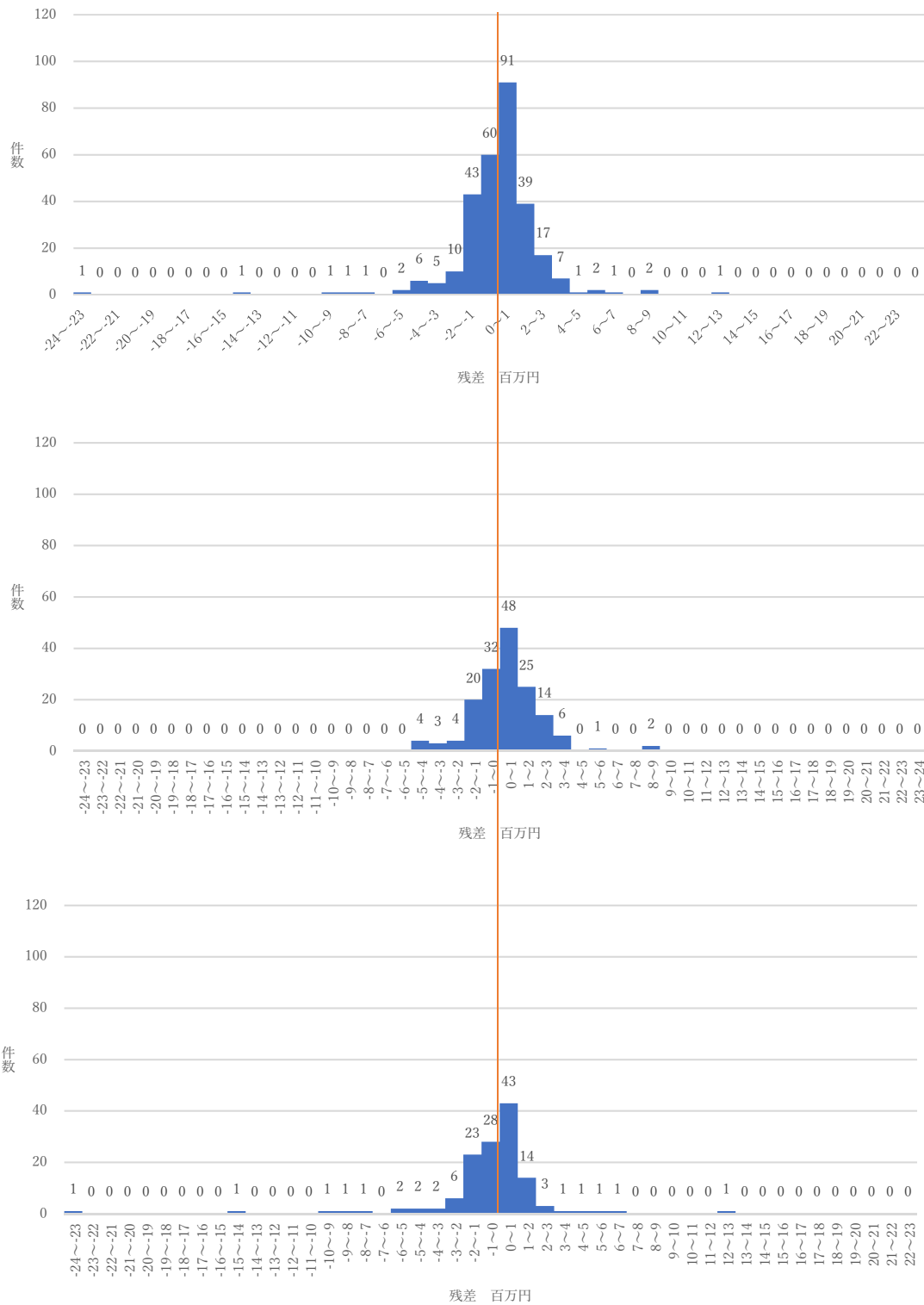


図9 OLS 回帰の残差の分布（応札者数の平均値を適用）
 （上：全局、中：北海道局、下：本州以南の局）

6. まとめ

分析及び検証を踏まえると、中央値回帰による算定式が最も適しているとの結果になった。実際に使用する場合の樹木料評定式の変数及び係数の組合せは別紙「樹木料評定式並びにその変数及び係数」のとおりである。樹木採取区には分収林が含まれないことから分収林ダミー変数は0となり、応札者数については今回使用したデータの平均値を適用することとする。年度ダミーについても、2015～2018年度の年度ダミー係数の平均値を用いることとする（これらの数値は別紙において定数に組み込まれている）。

なお、回帰式は誤差の範囲で算出額が変動することが想定される。そのため、実際の樹木料の算定に当たっては、運用のルールを設けることが望ましい。例えば、通常の立木販売の価格水準等を考慮し、その地域（販売ブロック）における過去1年間の立木販売実績単価を上限値として設けることなどが想定される。また、樹木採取権制度では、樹木採取権者は、毎年度、採取を行う箇所（伐区）を計画し、これを踏まえ、国が樹木料を算出して、樹木採取権者に通知する。樹木採取権者は、国から通知された伐区ごとの樹木料を確認した上で、実際に採取する箇所を選定できることとしている。このようなしくみを設けることで、樹木採取権制度における樹木料の算定方法が、樹木採取権者及び国の双方にとって納得の得られるしくみとなるものと期待される。

樹木料評定式並びにその変数及び係数

〈樹木料評定式〉

$$\log_e(\text{樹木料評定額}) = a_0 + \sum a_i x_i = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_{n-1} x_{n-1} + a_n x_n$$

北海道以外の局

分類	変数 x_i	係数 a_i		
量的データ	林分内容	面積 (10ha～) (ha)	-0.00583873	
		林齢 (年)	0.00303154	
		平均単木材積 (m3/本)	0.14407756	
		平均樹高 (m)	-0.00733571	
		複層伐材積比率	-0.40512687	
		間伐材積比率	-0.66256958	
		スギ材積比率	0.31813105	
		ヒノキ材積比率	0.50996361	
		カラマツ材積比率	0.94299672	
		広葉樹材積比率	-0.37119475	
	低質材材積比率	-0.20198344		
	丸太	log 丸太価額 (円)	1.71485660	
丸太単価 (円/m3)		-0.00005204		
経費	log 生産変動経費額 (円)	-0.63392982		
	log 生産固定経費額 (円)	-0.09665558		
カテゴリーデータ	販売ブロック	ブロック6_青森西部	-0.17983197	
		ブロック8_北岩手	0.07803766	
		ブロック9_南岩手	-0.28803013	
		ブロック10_宮城	0.14671565	
		ブロック11_秋田県北	-0.11196794	
		ブロック12_秋田県南	-0.07012493	
		ブロック13_山形	-0.11603241	
		ブロック14_浜通り	0.59941072	
		ブロック15_中通り	0.42232668	
		ブロック17_栃木	0.14907033	
		ブロック18_群馬	-0.08445743	
		ブロック19_新潟	-0.17969622	
		ブロック20_茨城	0.14824474	
		ブロック21_東京	-0.14282769	
		ブロック24_東北信	-0.25063248	
		ブロック28_愛知	0.19877571	
		ブロック29_紀伊	0.12983810	
		ブロック31_瀬戸内	0.05674899	
		ブロック33_四国	-0.13832024	
		ブロック34_北九州	0.04551394	
		ブロック35_熊本	0.05764722	
		ブロック36_大分	0.09307658	
		ブロック37_宮崎	0.07449192	
		定数	定数	-1.32647267

北海道局

分類	変数 x_i	係数 a_i	
量的データ	林分内容	立木材積 (m3)	-0.00009987
		面積 (ha)	-0.02893065
		面積 (10ha～) (ha)	0.02874778
		林齢 (年)	0.00183648
		平均単木材積 (m3/本)	-0.20610697
		本数密度 (本/ha)	-0.00014856
		複層伐材積比率	0.21467396
		カラマツ材積比率	0.48012594
		広葉樹材積比率	-0.70811763
		丸太	低質材材積比率
	原料材材積比率		2.18777780
	経費	log 丸太価額 (円)	3.06664510
		log 生産変動経費額 (円)	-1.60527520
		生産変動経費単価 (円)	-0.00004939
	カテゴリーデータ	販売ブロック	log 生産固定経費額 (円)
ブロック1_札幌			-0.09686672
		ブロック3_北見	0.15145865
定数	定数	-6.53129951	

参考資料

1. 分析に使用したデータ一覧
2. 統計用語の解説

分析に使用したデータ一覧

< Stepwise 法による分析の対象データ一覧 >

データ名称	step wise 採用	内容 (定義)
販売額	被説明変数	過去5年間(H26~30年度)の国有林における立木販売物件で、以下の条件で抽出された物件毎の販売価額(税抜総額) ①販売方法は一般競争に限定 ②混合契約(立木販売と造林事業の一括事業)は除外 ③官行造林(分収育林及び分収造林はダミー変数として分析に使用)は除外 ④施業方法は育成単層林、育成複層林に限定し、天然生林は除外 ⑤その他、支障木、風倒被害木等、樹木採取区として想定していないような販売物件は除外
立木材積	○	立木販売物件毎の立木材積
面積	○	立木販売物件毎の面積
面積(10ha~)	○	面積が10ha以上を超える場合、面積から10haを減じたもの。
ha蓄積(単位蓄積)	×	伐区の立木材積を伐区面積で除したもの。
林齢	○	立木販売物件の林齢 複数の林齢の樹木が混在している場合は、最も立木材積の多いの林齢を採用。
平均単木材積	○	立木材積を総本数で除したもの
本数密度	○	総本数を面積で除したもの
平均樹高	○	立木販売の対象木の樹高の総和を総本数で除したもの(本州以南の局のみ)
平均胸高直径	×	立木販売の対象木の胸高直径の総和を総本数で除したもの(本州以南の局のみ)
皆伐材積比率	-	立木材積に対する皆伐対象の立木材積の割合
複層伐材積比率	○	立木材積に対する複層伐対象の立木材積の割合
間伐材積比率	○	立木材積に対する間伐対象の立木材積の割合
択伐材積比率	○	立木材積に対する択伐対象の立木材積の割合
スギ材積比率	○	立木材積に対するスギの材積の割合
ヒノキ材積比率	○	立木材積に対するヒノキの材積の割合
カラマツ材積比率	○	立木材積に対するカラマツの材積の割合
トドエゾ材積比率	○	立木材積に対するトドマツ及びエゾマツの材積の割合
その他針葉樹材積比率	-	伐区の立木材積に対するその他針葉樹材積の割合
広葉樹材積比率	○	立木材積に対する広葉樹の材積の割合
低質材材積比率	○	伐区の立木材積に対する低質材の材積の割合
一般材(A材)材積比率	-	生産が見込まれる丸太材積に対する一般材の割合(北海道局のみ)
低質材(B材)材積比率	○	生産が見込まれる丸太材積に対する低質材の割合(北海道局のみ)

原料材（C材）材積比率	○	生産が見込まれる丸太材積に対する原料材の割合（北海道のみ）
丸太価額	○	販売物件毎の立木から生産される丸太の規格別の材積に、近隣の原木市場等の丸太価格（単価）を乗じた価額
丸太単価	○	丸太価額を丸太材積で除した価格 分析には12ヵ月移動平均、4ヵ月移動平均も使用したがStepwise法の過程で落とされた
生産変動経費	○	伐木造材、運搬、販売等に要する変動事業費
生産変動経費単価	○	生産変動経費を丸太材積で除したもの
生産固定経費	○	搬出施設費（林道改良費、修繕費を含む）、小屋掛費等で材積按分を適当としない固定事業費
生産固定変動費単価	×	生産固定経費を丸太材積で除したもの
応札者数	○	立木販売物件毎の応札者数（平均値2.89888153）
素材価格	×	木材価格統計の月次、地域別樹種別素材価格
分収育林、分収造林ダミー変数	○	直営林、分収育林、分収造林の別を表すダミー変数
年度ダミー変数	○	分析対象データの契約年度の別を表すダミー変数
季節ダミー変数	×	年度を3ヶ月毎に分けた四半期ごとのダミー変数
タイムトレンド（月次）	×	平成26年4月を1とし月ごとに1増えていく変数
販売ブロックダミー	○	国有林材の加工・流通圏域等を勘案して林野庁が設定した区域をダミー変数としたもの
素材価格	×	木材価格統計の月次、地域別樹種別素材価格

※「-」は、完全な多重共線性を防ぐために除外した変数

※「×」は、Stepwise法による分析の過程でP値が大きくなり除外されたもの

※「○」は、Stepwise法による分析の結果、残った変数

< Stepwise法による分析の対象から除外されたデータ一覧 >

データ名称	内容（定義）
予定価格	立木販売の物件毎に国が設定する最低価額
予定単価	予定価格を立木材積で除したもの
値開率	販売額を予定価格で除したもの
販売単価	販売額を立木材積で除したもの
地利	森林調査簿上の現在地利（0～24の地利指数）
林道距離	森林調査簿上の林道距離（区分は局によって異なる）
傾斜	森林調査簿上の傾斜（区分は局によって異なる）
標高	森林調査簿上の標高
局ダミー	販売を行った森林管理局を表すダミー変数

※データ収集したが、データに欠損がある場合や、検討委員会での議論を踏まえ不相当とされたもの等、最終的に分析には使用しなかったデータ

統計用語の解説

	用語	解説
エ	AIC (赤池情報量規 準)	統計的モデルのデータへの適合度の良さを評価する統計量。値 が小さいほど当てはまりが良い。
オ	折れ線回帰	回帰分析において、説明変数と被説明変数の関係が説明変数の 全域に渡って一直線では表せない場合に、説明変数がある点を 超えると回帰直線の傾きが変わることを許す定式化を行うこと があり、これを折れ線回帰や区分線形回帰などと呼ぶ。
カ	確率分布	変数が、ある値をとる確率を表したものの。
キ	疑似決定係数	中央値回帰分析などで、モデルのあてはまりを見る指標の一 つ。線形回帰分析の決定係数に相当することから、疑似決定係 数と呼ばれる。
	記述統計	データの分布を要約的に示す量のこと、平均や中央値、分 散、最大値、最小値などがよく用いられる。
ク	クロスバリデーション	データを分割し、その一部でまずモデルを推定して、残る部分 でその推定したモデルのテストを行い、推定の妥当性の検証・ 確認を行う手法のこと。
ケ	決定係数 (R^2)	回帰分析において、被説明変数の実現値に対する非説明変数の 予測値の説明力を表す指標。0 から 1 までの値をとり、1 に近 いほど推定された回帰式が有効である。
サ	最小二乗法	観測値と予測値の誤差の二乗和が最小になるように予測値のパ ラメーターを求める方法。
	残差	回帰分析において、推定された回帰式から得られる被説明変数 の予測値から実現値を引いたもの。
	残差平方和	残差の平方(2乗)を標本全体で足し合わせたもの。
シ	重回帰分析	多変量解析の中で最も多用される分析手法で、被説明変数と複 数の説明変数との関係を分析する方法。
ス	Stepwise 変数減少法	重回帰分析において、一定の基準を用いて説明力の低い説明変 数を取り除きながら効果的な説明変数だけで構成されるモデル を選択する方法。
セ	説明変数	多変量解析において原因となる変数のこと。
	絶対残差	残差の絶対値のこと。
タ	対数変換	変数の対数を取ること。すなわち、変数 X に対して新たな変 数 $\ln(X)$ を作ること。

	用語	解説
	多重共線性	重回帰分析において、2つ以上の説明変数、例えば X1 と X2 の間で相関関係が強いために、被説明変数 Y への X1 の影響と X2 の影響を分離し難いという推定上の問題が生じること。
	多変量解析	3つ以上の変数を同時に扱う分析手法の総称、予測や変数間の因果関係の解析などに用いられる。
	ダミー変数	量的データではない質的データのうちカテゴリーを表す変数として表したものの。具体的には、あるカテゴリーに当てはまる場合に 1 を、当てはまらない場合に 0 を取る変数。
チ	中央値	ある変数のデータを値の小さいほうから順に並べた時、ちょうどデータの半分に位置する値のこと。平均とは異なる概念だが、やはりその変数の分布の中心を表すもの。
	中間値回帰	重回帰分析の一手法で、説明変数の取る値が決まった時の被説明変数の中央値を予測する式を推定する方法。
ハ	外れ値	全体の傾向とは異なって、離れた値を示すデータのこと。ヒストグラムや散布図で検出することができる。
ヒ	ヒストグラム	データの分布を見るため、各階級に属するデータの数を棒グラフで図化して表したものの。
	被説明変数	多変量解析において結果となる変数のこと。
	p 値	統計的検定で用いられる指標。ある仮説、典型的には回帰式の回帰係数が 0 であるとの仮説を立てた場合に、その仮説の確からしさを表現するもの。0 から 1 の間の値を取り、0.05(5%) 未満であれば、その係数は統計的有意であると言う。
フ	V I F (分散拡大要因)	説明変数間の多重共線性を検出するための指標の 1 つ。値が 10 を超える場合は、多重共線性は発生していると判断される場合が多い。
ヘ	平方残差	残差の平方 (二乗) のこと。
ユ	尤度	Y を予測する回帰式 (確率モデル) を考えた場合、式の係数を特定した時に Y が観測された値を取る確率のこと。標本全体の尤度が最大になるように係数を決定する最尤法は回帰式推定の代表的な方法で、誤差項に正規分布を仮定する場合、最尤法は OLS と一致する。
ロ	ロバスト推定	重回帰分析では、回帰式の係数を推定するとともに、係数推定値の分散の大きさを推定する。その際に、モデルに強い仮定をおかずに推定を行う方法。