

住宅エネルギー消費の世帯間における差異と その発生要因に関する分析

～政府統計データに基づく、機械学習によるアプローチ～

国立研究開発法人 建築研究所

環境研究グループ

羽原 宏美

本日の内容

はじめに

- 研究背景
- 研究目的

データおよび 分析フロー

- 使用データ：家庭部門のCO2排出実態統計調査
- 補 足：一次エネルギー・二次エネルギー
- 分析用データ：年間一次エネルギー消費量・用途別構成比率
世帯類型の定義
特徴量（121変数）の一覧
- 分析の全体フロー

世帯間における 差異に関する分析

- 補 足：地域の区分
- 分析結果：年間一次エネルギー消費量
用途構成（用途別構成比率の代表値の推定）

差異の発生要因に 関する分析

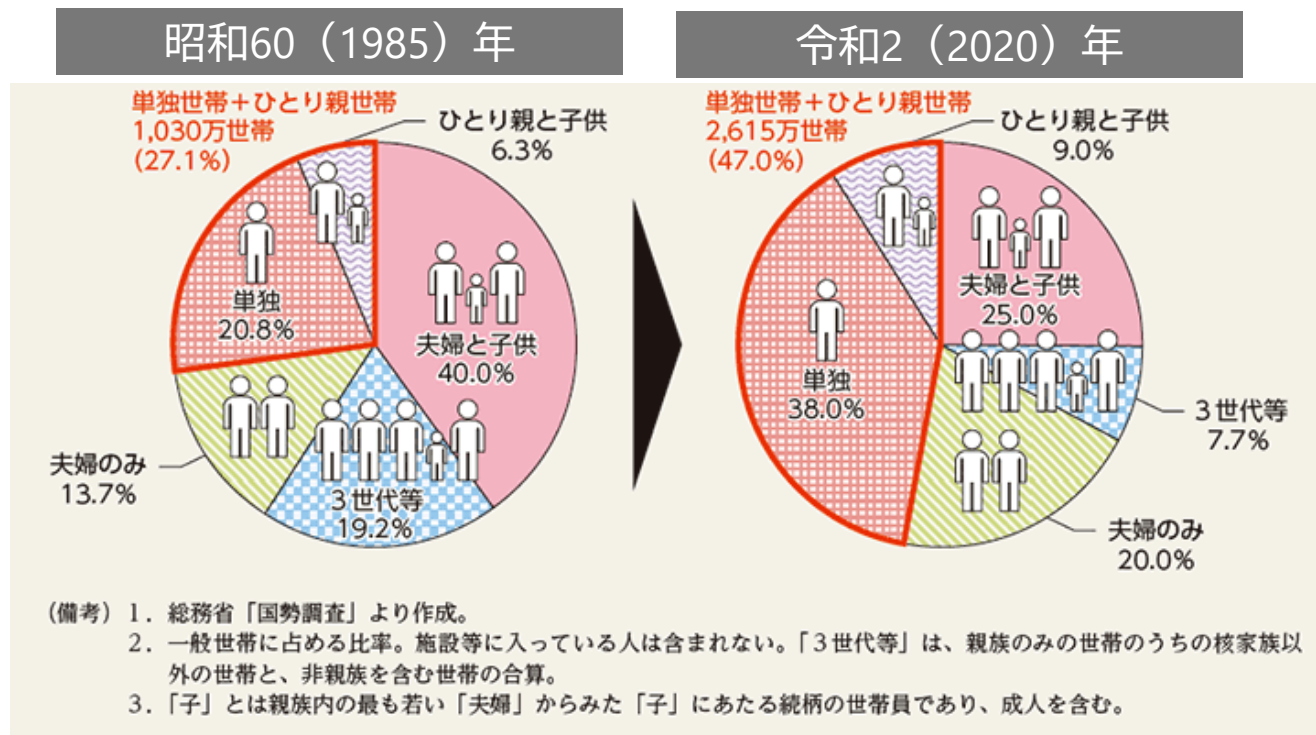
- 分析方法：機械学習による分析の流れ
- 補 足：機械学習モデルの解釈手法SHAPについて
- 分析結果：特徴量重要度（上位10変数）
SHAP値の分布（延床面積）
要因分解

まとめ

はじめに

■ 研究背景

- 少子高齢化等に伴う世帯構成・規模の変化は、住宅におけるエネルギーの使い方に影響を与えている。
- 今後さらなる住宅の性能向上に向けた対策を考える上では新たな視点が必要となることが予想される。



【出处】 内閣府男女共同参画局：男女共同参画白書 令和5年版

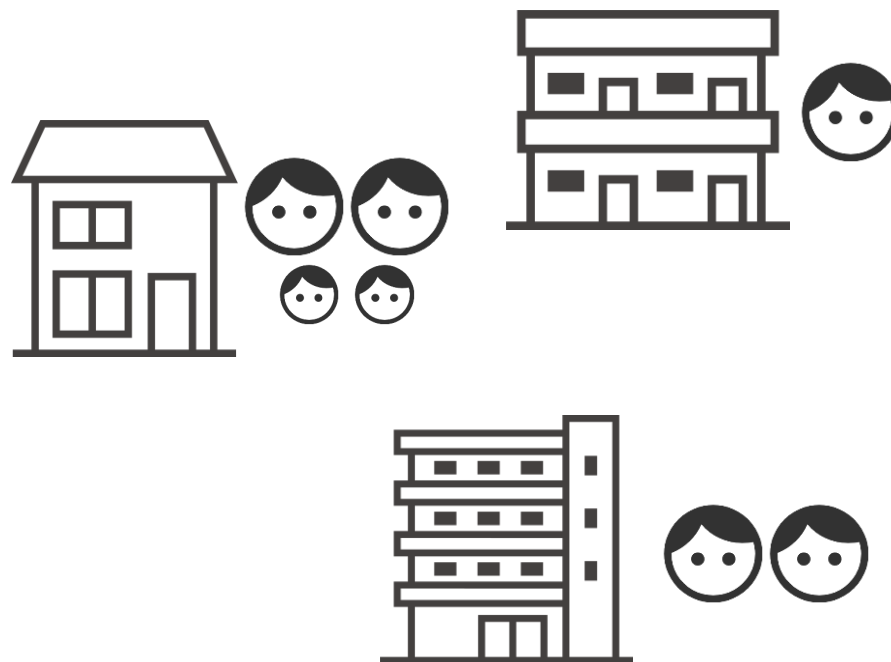
(https://www.gender.go.jp/about_danjo/whitepaper/r05/zentai/html/zuhyo/zuhyo00-op01.html)

はじめに

■ 研究目的

- 環境省が実施する「家庭部門のCO2排出実態統計調査（家庭CO2統計）」の調査票情報に基づいて住宅におけるエネルギー消費の実態を分析し、次の事項を把握する。
 - ・ 世帯間における差異
 - ・ 差異の発生要因

どのくらい 違うの？
なぜ 違いが生まれるの？



データおよび分析フロー

■ 使用データ：家庭部門のCO₂排出実態統計調査（家庭CO₂統計）

- 家庭CO₂統計は、全国の世帯を対象とした**政府の一般統計調査**であり、家庭からのCO₂排出量やエネルギー消費量の実態を把握することを目的としている。
- 本研究では、環境省に対して統計法第33条に基づいて二次的利用の申請を行い、**6カ年度分**（平成29年度～令和4年度）の調査票情報の提供を受けた。

The screenshot shows the Japanese Ministry of the Environment website. At the top left is the logo and name '環境省 Ministry of the Environment'. There are navigation links for '本文 >' and 'English'. A search bar contains 'キーワード検索' and a '検索' button. Below the search bar is a green banner with the text '地球環境・国際環境協力'. Underneath is a breadcrumb trail: 'ホーム > 政策 > 政策分野一覧 > 地球環境・国際環境協力 > 地球温暖化対策 > 家庭部門のCO₂排出実態統計調査（家庭CO₂統計）'. The main heading is '家庭部門のCO₂排出実態統計調査（家庭CO₂統計）'. The main text explains that the survey aims to grasp CO₂ emissions and energy consumption in households across the country. It mentions the survey is based on the Statistics Act (Act No. 53 of 1998). A note at the bottom states that the survey is conducted for households that have responded to the survey.

調査の地域	全国
調査の対象	13,000世帯（店舗併用住宅等を除く）
調査期間	調査実施年の4月～翌年3月（毎月）
調査事項	<ol style="list-style-type: none">① 月別のCO₂排出量を推計するためのエネルギー使用量等について② 太陽光発電について③ 設備・機器について④ 世帯について⑤ 住宅について⑥ 家電製品等について⑦ 給湯について⑧ コンロ・調理について⑨ 車両について⑩ 暖房機器について⑪ 省エネルギー行動の実施理由

【出处】 環境省HP (<https://www.env.go.jp/earth/ondanka/ghg/kateiCO2tokei.html>)

データおよび分析フロー

■ 補足：一次エネルギー・二次エネルギー

- 一次エネルギー：自然界から直接採取できるエネルギー源。
- 二次エネルギー：一次エネルギーを加工・変換して得られるエネルギー。

一次エネルギー

変換・加工

二次エネルギー

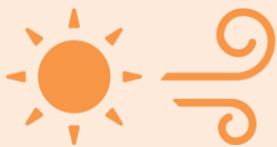
用途



石油・天然ガス



石炭



太陽光・風



石油・ガス精製所



火力発電所



太陽光・
風力発電所



灯油



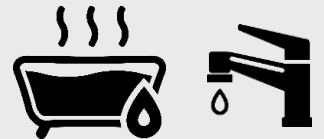
都市ガス・LPガス



電気



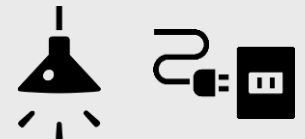
暖房・冷房



給湯



調理



照明他

データおよび分析フロー

■ 分析用データ：年間一次エネルギー消費量・用途別構成比率

- 家庭CO2統計の調査票情報から電気・ガス・灯油の月別使用量および太陽光発電の月別発電量・売電量を取得。
- 家庭CO2統計の推計手順に準ずる方法により、上記の月別使用量等を用途別の電気・ガス・灯油の消費量に分解。
- 各種換算係数を用いて、用途別消費量から一次エネルギー消費量と用途別構成比率を算定。



データおよび分析フロー

■ 分析用データ：世帯類型の定義



本分析における区分名	世帯構成	世帯員の年代
単 独 (非高齡)	世帯主のみ	65歳未満
単 独 (高 齡)		65歳以上
夫婦のみ (非高齡)	世帯主 + 世帯主の配偶者	世帯主または世帯主の配偶者のいずれか、もしくはその両方の年齢が65歳未満
夫婦のみ (高 齡)		世帯主または世帯主の配偶者の両方の年齢が65歳以上
夫婦と子 (未成年)	世帯主 + 世帯主の配偶者 + 子・子の配偶者	子・子の配偶者の全員の年齢が20歳未満
夫婦と子 (成 年)		子・子の配偶者に年齢が20歳以上である者を1人以上含む
その他	上記以外	区別なし

※一部の分析においては、夫婦と子世帯を
居住人数（3人、4人、5人以上）によりさらに分類。

データおよび分析フロー

■ 分析用データ：特徴量（121変数）の一覧

要因		特徴量（説明変数）	要因		特徴量（説明変数）
大分類	小分類		大分類	小分類	
地域	都市階級	都市階級	給湯	使用有無	電気ヒートポンプ式給湯機／電気温水器／ガス給湯機・ガス風呂がま／ガス小型瞬間湯沸器／灯油給湯機・灯油風呂がま／太陽熱利用給湯機／ガスエンジン発電・給湯機／家庭用燃料電池／その他給湯機の使用有無
	気候	暖房度日（HDD18-18）[度日]、冷房度日（CDD24-24）[度日]、給湯度日（WDD40-40）[度日]、年積算全天日射量[MJ/m ²]		使い方	夏季湯はり割合[-]、夏季シャワー割合[-]、夏季入浴なし割合[-]、冬季湯はり割合[-]、冬季シャワー割合[-]、冬季入浴なし割合[-]、洗面所のお湯の使用頻度、台所のお湯の使用頻度
世帯	世帯構成	居住人数[人]、世帯類型	調理	使用有無	ガスコンロ／電気コンロ／その他のコンロの使用有無
	経済状況	世帯収入、就業割合[-]		使い方	1週間当たりの調理食数[食]
	住まい方	平日在宅	テレビ	使用台数	テレビの使用台数[台]
建て方	建て方	種類		1/2/3台目テレビの種類	
建築時期	建築時期	画面サイズ		1/2/3台目テレビの画面サイズ	
所有関係	所有関係	製造時期		1/2/3台目テレビの製造時期	
住宅規模	延床面積[m ²]、居室数[室]	使い方		テレビの平日使用時間	
暖房	窓仕様	二重サッシまたは複層ガラスの有無	冷蔵庫	使用台数	冷蔵庫の使用台数[台]
	使用有無（セントラル暖房／床暖房／太陽熱）	セントラル暖房のエネルギー源		種類	1/2台目冷蔵庫の種類
		電気床暖房／ガス温水床暖房／灯油温水床暖房の使用有無		庫内容積	1/2台目冷蔵庫の庫内容積
		太陽熱利用暖房システムの使用有無		製造時期	1/2台目冷蔵庫の製造時期
冷房	使用台数（その他）	エアコン／電気ストーブ類／電気カーペット・こたつ／電気蓄熱暖房器／ガスストーブ類／灯油ストーブ類／木質系燃料ストーブ類の使用台数[台]	洗濯乾燥	使用台数	洗濯機（乾燥機能なし）／洗濯機（乾燥機能付）／衣類乾燥機（電気）／衣類乾燥機（ガス）／浴室乾燥機の使用台数[台]
	使い方	暖房の仕方、主暖房機器、主暖房機器の設定温度[°C]、主暖房機器の強弱設定、暖房の平日使用時間、暖房居室数、ペットのための暖房使用		使い方	乾燥機能使用頻度
	使用台数	冷房機器の使用台数[台]	トイレ便座	使用台数	温水洗浄便座／暖房便座（温水洗浄機能なし）の使用台数[台]
	種類	1/2/3/4/5台目エアコンの種類		ペット用品	使用有無
製造時期	1/2/3/4/5台目エアコンの製造時期	太陽光	容量		太陽光発電の容量[kW]
使い方	冷房の平日使用時間、冷房の設定温度[°C]、ペットのための冷房使用		省エネ意識	使い方	省エネ行動*の実施割合[-] （※照明：調光/消灯、給湯：シャワー・入浴時・食器洗い時の工夫/コントロールのオフ、調理：電子レンジで下ごしらえ/火力調整、テレビ：明るさ調整/主電源オフ、冷蔵庫：温度調整、つめこみ過ぎ、トイレ便座：温水の設定温度の調整、冬季以外暖房便座機能オフ）
照明	使用有無（場所ごと）	居間／食卓／台所／個室／その他の場所の白熱電球の使用有無、居間／食卓／台所／個室／その他の場所の蛍光灯の使用有無、居間／食卓／台所／個室／その他の場所のLEDの使用有無、居間／食卓／台所／個室／その他の場所のその他のランプの使用有無			
	使い方	主照明の種類、居間照明の使用時間			

※ 目的変数（エネルギー消費量に関連する変数）を除く

データおよび分析フロー

分析の全体フロー

事前準備

データ前処理・特徴量データの作成と分析対象世帯の選定

使用データ：調査票情報

- エネルギーに関する事項
 - ・ 電気・ガス・灯油の月別使用量
 - ・ 太陽光発電の月別発電量・売電量
- 地域に関する事項
- 世帯に関する事項
- 住宅建物に関する事項
- 使用機器に関する事項

データ前処理

- 以下の事項に起因する実施年度や設問によるデータの差異を解消。
 - ・ 実施年度による調査票の相違
 - ・ 同種の設問間における回答選択肢の表記ゆれ
 - ・ 回答『不明』の扱い

特徴量データの作成

- 換算係数により年間一次エネルギー消費量を算定。
- 調査票情報に基づき、用途別エネルギー消費量を算定。
- 市区町村役場の緯度経度に基づき、「ArcClimate」「grib2-to-csv」により各地点の気象データを作成。
- その他、適宜、データ変換・整形。

特徴量の数：121 変数
(説明変数のみ)

分析対象世帯の選定

- 【選定条件】
- 家庭CO2統計において用途別エネルギー消費量の推計対象である。
 - 用途別エネルギー消費量の算定に必要な情報が全て揃っている。
 - 実績値に対する算定値の乖離が10%以内である。
 - 算定値が下限条件を満たしている。

分析対象：42,370 サンプル

統計分析

世帯間における差異に関する分析

- 世帯を世帯類型・居住人数に基づいて分類。
- 地域の区分ごとに層別し、以下の分析を実施。
 - ・ 年間一次エネルギー消費量・用途別構成比率に関する群間差の検定
 - ・ 用途別構成比率の代表値の推定

機械学習

差異の発生要因に関する分析

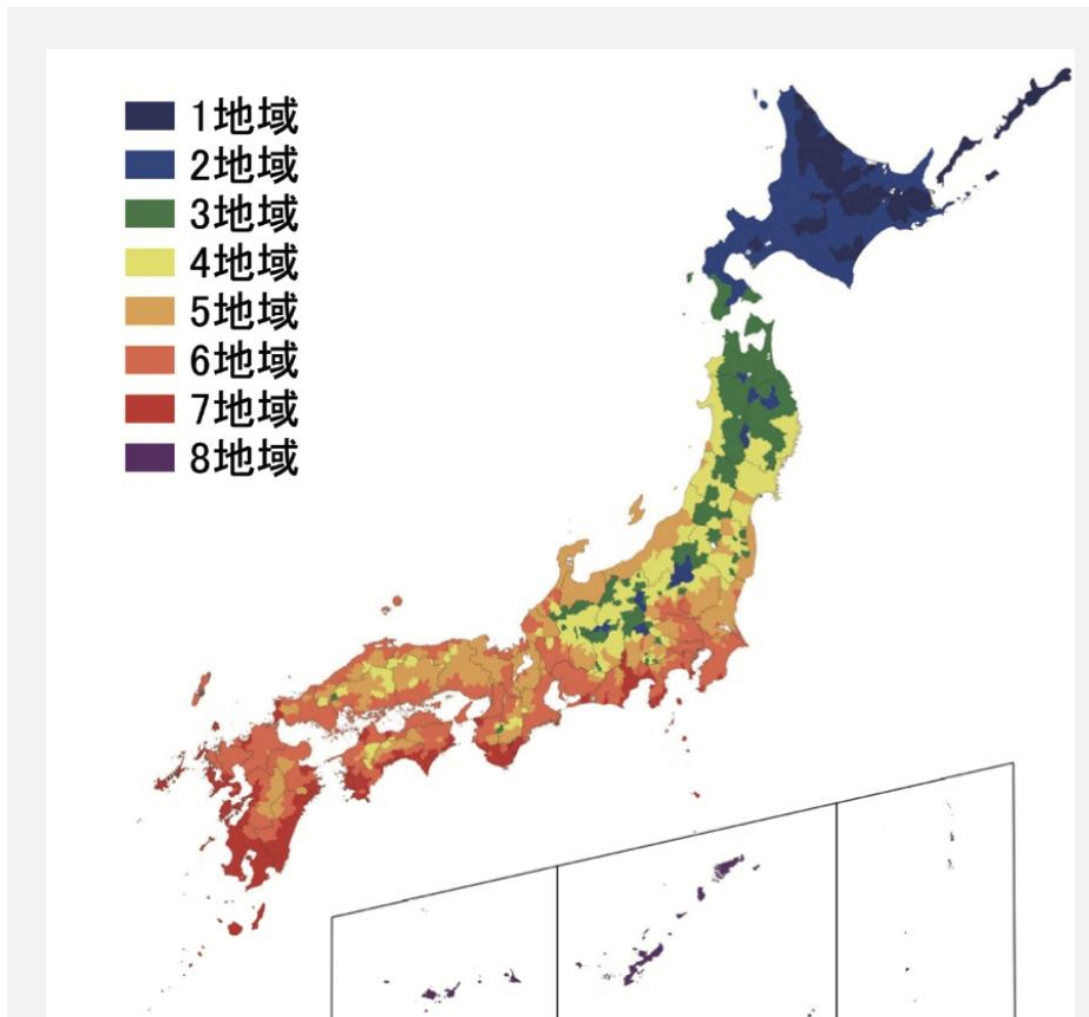
- 機械学習により年間一次エネルギー消費量・用途別構成比率に関する予測モデルを作成。
- 作成した予測モデルの解釈を通じて、年間一次エネルギー消費量・用途別構成比率に世帯により差異が生じる要因を分析。

世帯間における差異に関する分析

■ 補足：地域の区分

- 「地域の区分」とは、省エネルギー基準（国土交通省告示第265号）が定める地域の区分であり、気候により全国を8地域（1地域～8地域）に区分される。
- 省エネルギー基準では、外皮性能やエネルギー消費性能の基準値を地域の区分ごとに定めている。

※ 以降においては、
東京・大阪を含む「6地域」
を中心に分析結果を示す。



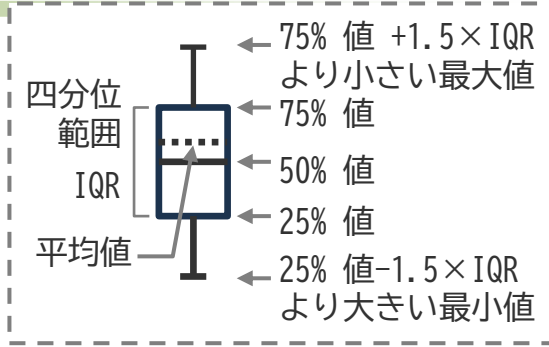
【出处】

国土交通省HP (<https://www.mlit.go.jp/shoene-label/insulation.html>)

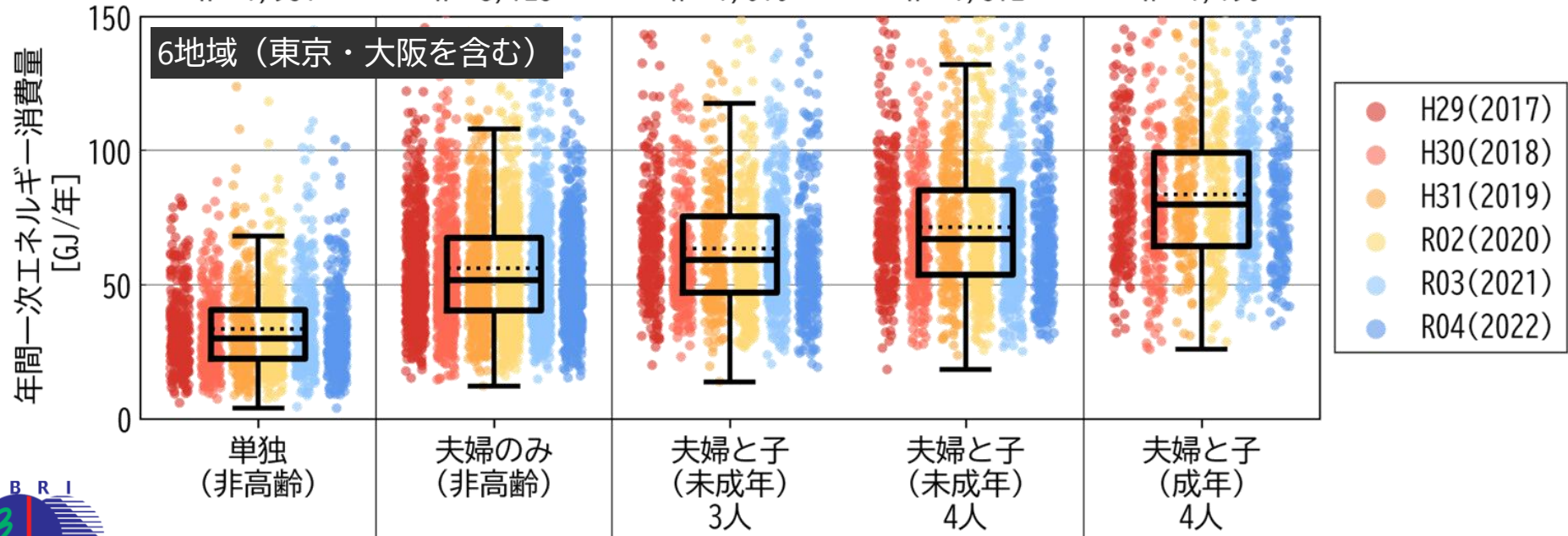
世帯間における差異に関する分析

■ 分析結果：年間一次エネルギー消費量

- 平均値は、夫婦と子（成年）世帯、夫婦と子（未成年）世帯、夫婦のみ世帯、単独世帯の順に大きかった。
- ばらつきは、居住人数が多いほど大きかった。



平均	33.41 GJ	<	56.08 GJ	<	63.39 GJ	<	71.46 GJ	<	83.61 GJ
標準偏差	16.38 GJ	<	23.17 GJ	<	23.21 GJ	<	25.49 GJ	<	28.00 GJ
	n = 1,951		n = 3,723		n = 1,610		n = 1,872		n = 1,190



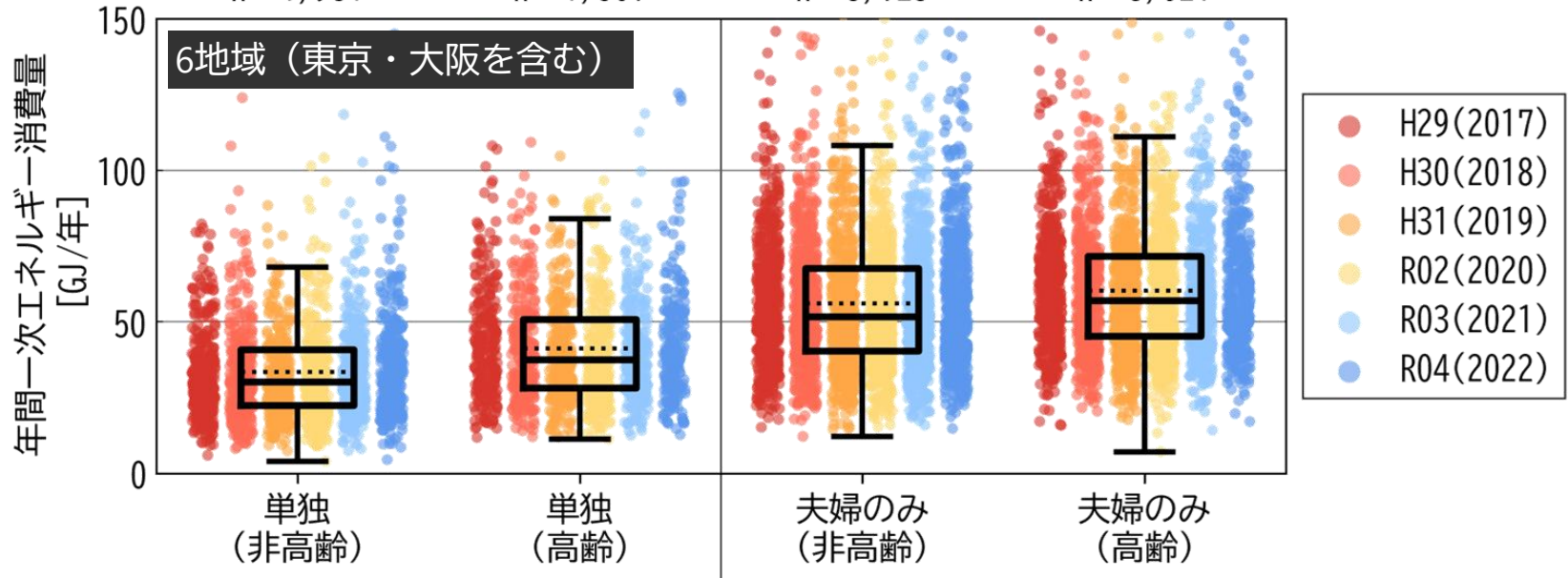
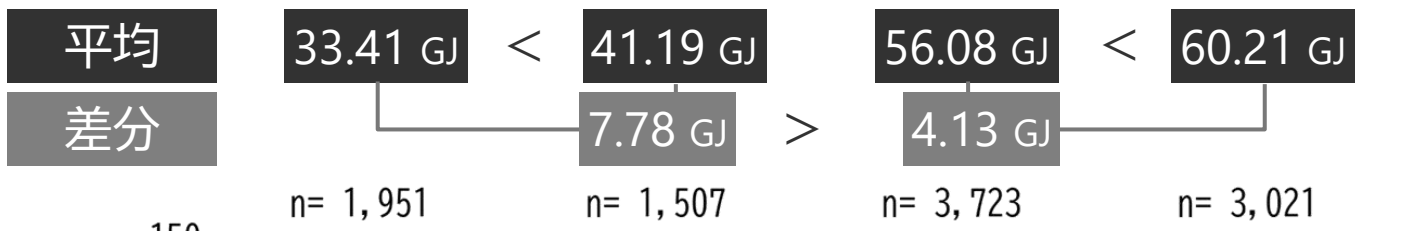
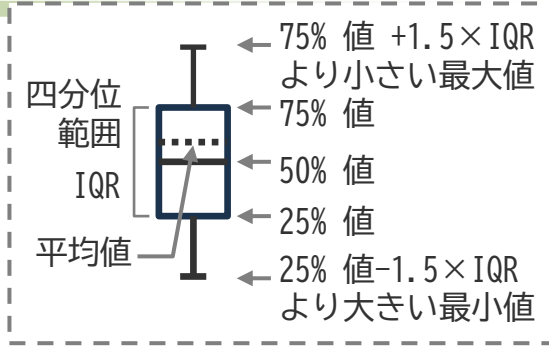
多重比較の結果、すべての群間の組み合わせで有意差あり (p < 0.05)



世帯間における差異に関する分析

■ 分析結果：年間一次エネルギー消費量

- 平均値は、高齢世帯の方が非高齢世帯よりも大きかった。
- 非高齢世帯と高齢世帯の平均の差は、単独世帯の方が大きかった。



多重比較の結果、すべての群間の組み合わせで有意差あり (p < 0.05)



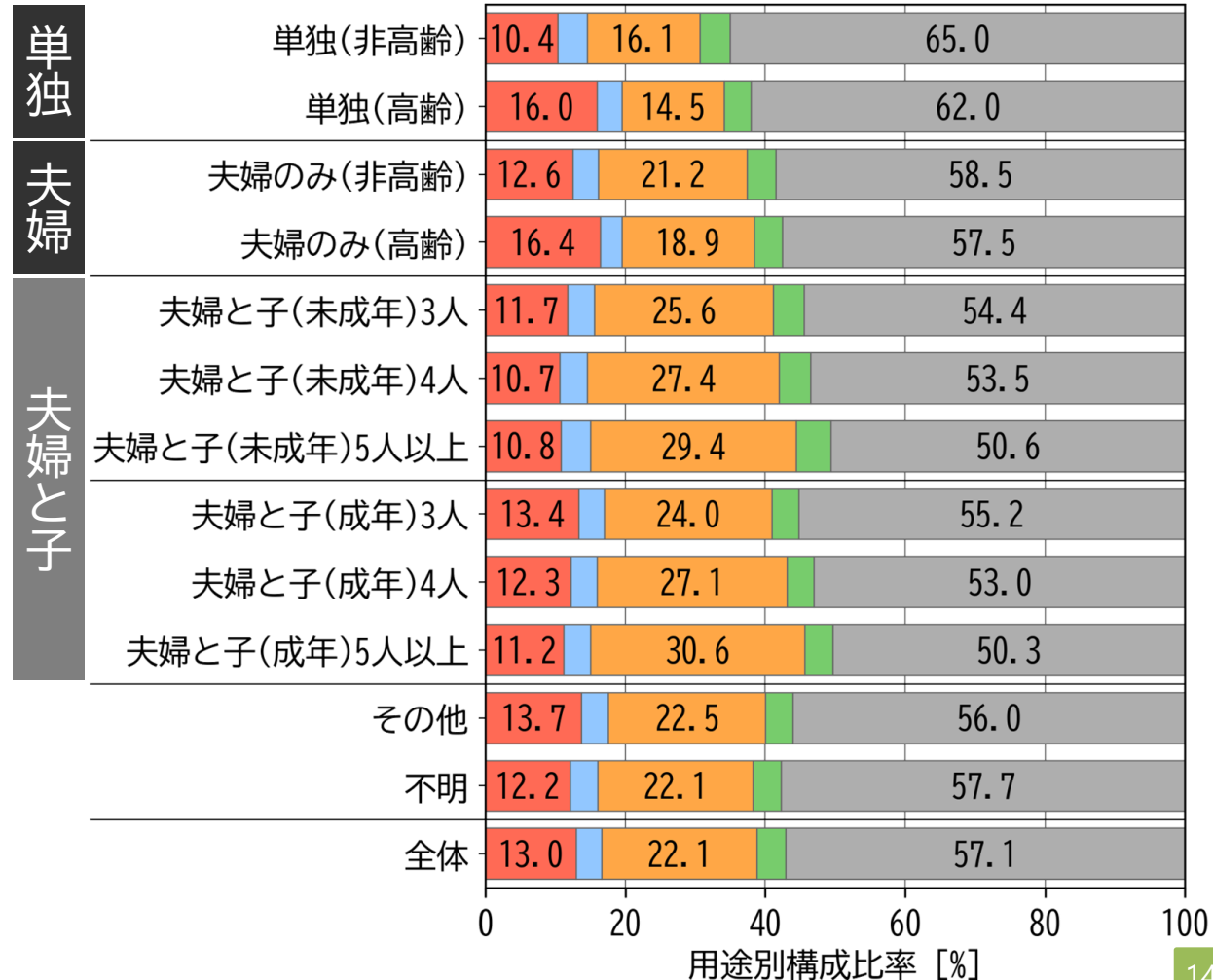
世帯間における差異に関する分析

■ 分析結果：用途構成（用途別構成比率の代表値の推定）

- 全体としては、居住人数が多いほど、給湯・調理・照明他の比率が高い傾向がみられた。
- 単独世帯および夫婦のみ世帯では、高齢世帯は非高齢世帯に比べて暖房比率が高く、給湯比率が低かった。
- 夫婦と子世帯では、人数差による影響は限定的であった。

■ 暖房 ■ 冷房 ■ 給湯 ■ 調理 ■ 照明他

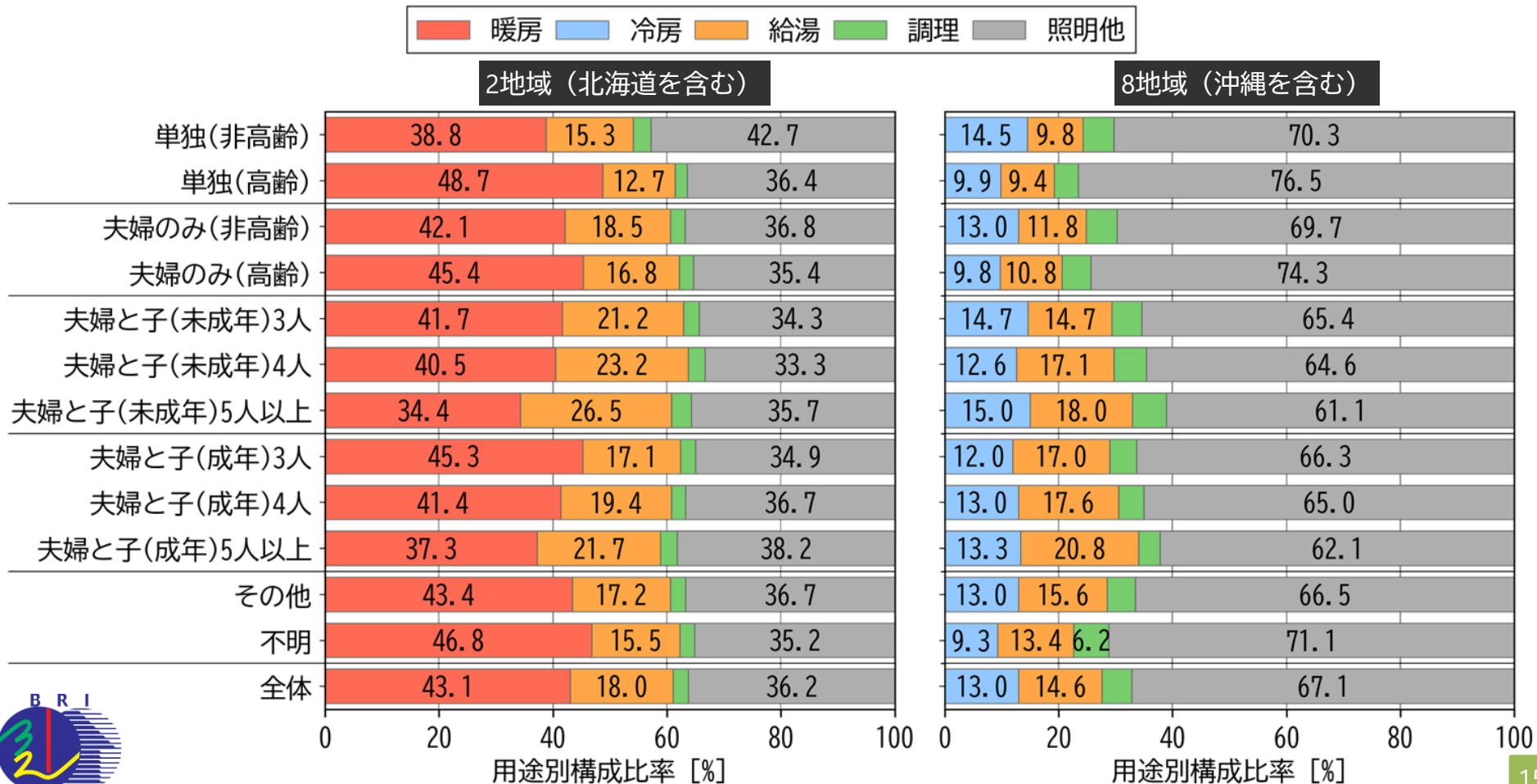
6地域（東京・大阪を含む）



世帯間における差異に関する分析

【参考】用途構成（用途別構成比率の代表値の推定）

- 1地域～8地域の全ての地域の区分について、用途別構成比率の代表値を推定。
- 地域の区分や世帯類型・居住人数による用途構成の違いを分析。



差異の発生要因に関する分析

■ 分析方法：機械学習による分析の流れ

事前準備

データ前処理・特徴量データの作成と分析対象世帯の選定

使用データ：調査票情報

データ前処理

特徴量データの作成

分析対象世帯の選定

特徴量の数：121 変数
(説明変数のみ)

分析対象：42,370 サンプル

モデル作成

勾配ブースティング決定木による機械学習モデルの作成

アルゴリズム

XGBoost
(eXtreme Gradient Boosting)

ハイパーパラメータ探索方法

ランダムサーチ

モデル評価

4 分割
交差検証

評価指標

RMSE

予測

機械学習モデルによる全分析対象世帯の年間一次エネ消費量※の予測

※ 用途構成の分析においては、用途別構成比率

モデル解釈

SHAPによる予測における差異の発生要因の分析

特徴量重要度

- SHAP値の絶対値を特徴量ごとに平均。
- 予測における特徴量の平均的な寄与度を分析。

SHAP値の分布

- 世帯ごと・特徴量ごとのSHAP値を使用。
- 予測における各特徴量の寄与の程度と方向を分析。

要因分解

- SHAP値を要因ごとに合計し、これを世帯群ごとに平均。
- 予測値の構造をマクロに分析。

差異の発生要因に関する分析

■ 補足：機械学習モデルの解釈手法SHAPについて

- SHAP (SHapley Additive exPlanations) は、協力ゲーム理論のShapley 値を応用し、機械学習モデルの予測値に対する特徴量 (説明変数) の寄与度を計算する手法。
- 協力ゲーム理論のShapley 値は、協力ゲームで得られる報酬をプレイヤーの貢献度に応じて分配するための方法の一つ。

<例> Aさん、Bさん、Cさんが協力してゲームに挑戦して、賞金60万円を得た。3人にどのように分配すればよいか？

				限界貢献度 [万円]		
				Aさん	Bさん	Cさん
なし	+ Aさん	+ Bさん	+ Cさん	10	20	30
0万円	10万円	30万円	60万円	(= 10 - 0)	(= 30 - 10)	(= 60 - 30)
なし	+ Aさん	+ Cさん	+ Bさん	10	38	12
0万円	10万円	22万円	60万円	(= 10 - 0)	(= 60 - 22)	(= 22 - 10)
なし	+ Bさん	+ Aさん	+ Cさん	24	6	30
0万円	6万円	30万円	60万円	(= 30 - 6)	(= 6 - 0)	(= 60 - 30)
なし	+ Bさん	+ Cさん	+ Aさん	44	6	10
0万円	6万円	16万円	60万円	(= 60 - 16)	(= 6 - 0)	(= 16 - 6)
なし	+ Cさん	+ Aさん	+ Bさん	18	38	4
0万円	4万円	22万円	60万円	(= 22 - 4)	(= 60 - 22)	(= 4 - 0)
なし	+ Cさん	+ Bさん	+ Aさん	44	12	4
0万円	4万円	16万円	60万円	(= 60 - 16)	(= 16 - 4)	(= 4 - 0)
Shapley値：限界貢献度の平均値				25万円	20万円	15万円

【出处】DataRobot社：
「SHAP を用いて
機械学習モデルを説明する」

<https://www.datarobot.com/jp/blog/explain-machine-learning-models-using-shap/>

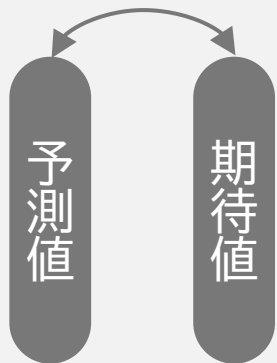
差異の発生要因に関する分析

■ 補足：機械学習モデルの解釈手法SHAPについて（続き）

- SHAP 値は、あるインスタンス（ここでは「世帯」）において予測値と期待値との間に生じた差分を各特徴量の寄与度に分解したものである。
- SHAP は「機械学習モデルがなぜそのような予測を出したのか」を解釈する手法である。（注意：因果関係を説明するものではない）。

差分 ● 各特徴量の寄与度（SHAP値）に分解

※ 「報酬の総額」が「予測値と期待値の差分」に、「プレイヤー」が「説明変数」にそれぞれ対応する。



各特徴量のSHAP値

世帯	ID	予測値	-	期待値	=	各特徴量のSHAP値								
						都市階級	暖房度日	...	太陽光発電の容量	省エネ行動の実施割合				
	ID:00001	53.46	-	68.24	=	0.07	+	2.53	+	...	+	-0.08	+	1.47
	ID:00002	38.65	-		=	0.02	+	3.26	+	...	+	-0.09	+	-0.47
	-		=	...	+	...	+	...	+	...	+	...
	ID:42370	58.06	-		=	0.00	+	-0.47	+	...	+	-0.06	+	-3.16

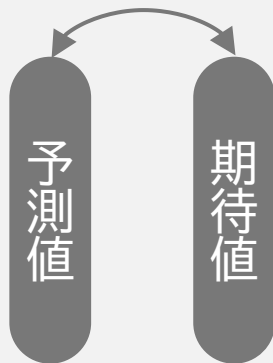
差異の発生要因に関する分析

■ 補足：機械学習モデルの解釈手法SHAPについて（続き）

- SHAP値には**加法性**があり、特定の属性群に属するサンプルに限定して集計することで、群間における寄与構造の差異を分析することが可能である。
- 特に、特徴量ごとにSHAP値の絶対値を平均した指標は「**特徴量重要度**」と呼ばれ、予測における各特徴量の平均的な寄与の大きさを表す指標として広く用いられる。

差分 ● 各特徴量の寄与度（SHAP値）に分解

※ 「報酬の総額」が「予測値と期待値の差分」に、「プレイヤー」が「説明変数」にそれぞれ対応する。



各特徴量のSHAP値

ID	報酬	予測値	期待値	各特徴量のSHAP値									
				都市階級	暖房度日	...	太陽光発電の容量	省エネ行動の実施割合					
ID:00001	53.46	-	68.24	=	0.07	+	2.53	+	...	+	-0.08	+	1.47
ID:00002	38.65	-		=	0.02	+	3.26	+	...	+	-0.09	+	-0.47
...	...	-		=	...	+	...	+	...	+	...	+	...
ID:42370	58.06	-		=	0.00	+	-0.47	+	...	+	-0.06	+	-3.16

差異の発生要因に関する分析

■ 分析結果：特徴量重要度（上位10変数）

SHAP値集計：特徴量ごとに
絶対値を平均
分析内容：平均的な寄与度

- 寄与度が大きい上位2変数は、「居住人数」「延床面積」であった。
- 上記の他、暖房・給湯に関連する特徴量が多く挙がった。

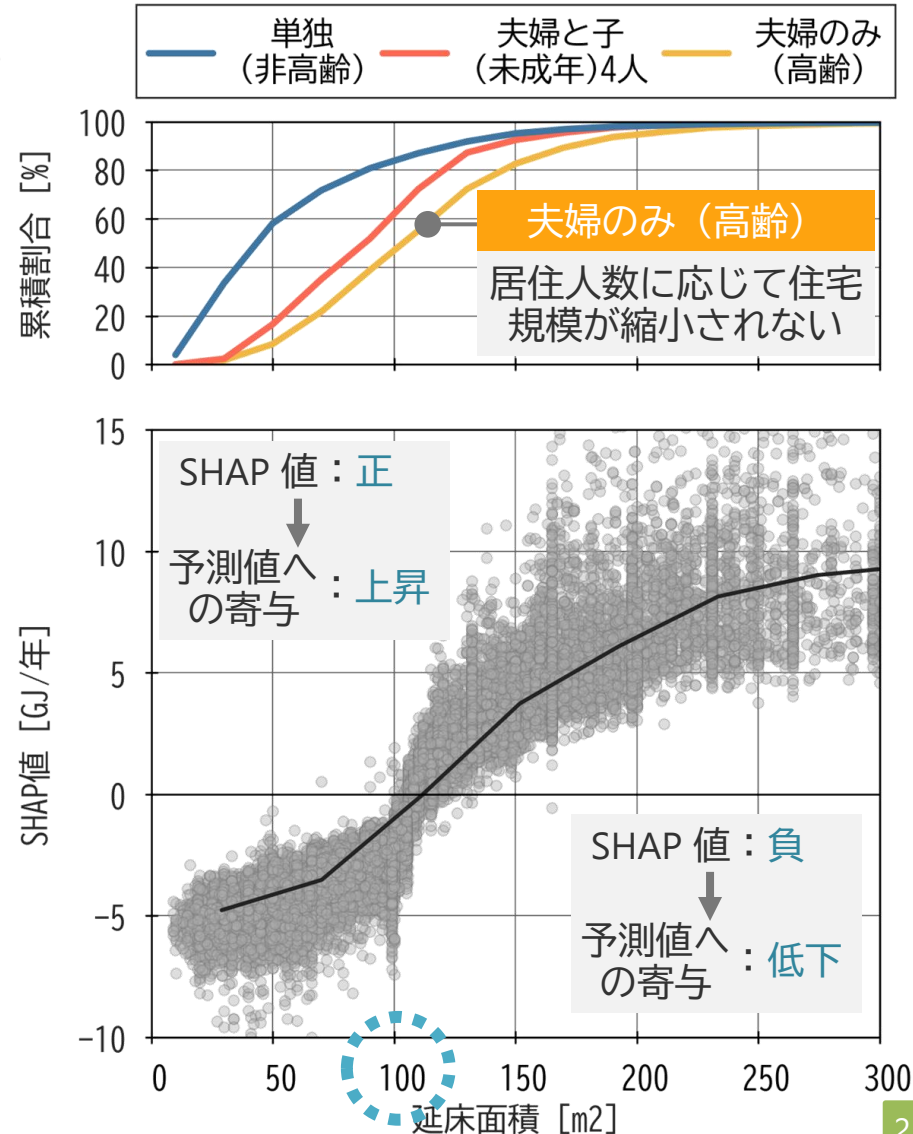
ランク	特徴量	区分	特徴量重要度
1	居住人数	世帯	6.83
2	延床面積	住宅建物	3.58
3	電気温水器の使用有無	給湯	2.86
4	暖房の仕方	暖房	2.67
5	ガス給湯器・ガス風呂がまの使用有無	給湯	2.61
6	暖房居室数	暖房	2.44
7	主暖房機器	暖房	2.16
8	暖房の平日使用時間	暖房	2.13
9	1週間当たりの調理食数	調理	1.84
10	省エネ行動の実施割合	省エネ意識	1.66
次点	暖房度日	地域	1.61

差異の発生要因に関する分析

SHAP値集計：なし
分析内容：寄与の程度と方向

■ 分析結果：SHAP値の分布（延床面積）

- 100m² 付近を境に SHAP 値の符号が反転。
 - 寄与の方向：
 - 100m²未満：予測値を低下させる
 - 100m²以上：予測値を上昇させる
- 夫婦のみ（高齢）では、居住人数に応じて住宅規模が縮小されない傾向が見られる。多くの世帯が SHAP 値が正となる領域にあり、延床面積は予測値を押し上げる要因として作用する。



世帯群の区分	SHAP値 [GJ/年]		
	平均	信頼区間	
		下側	上側
単独（非高齢）	-3.01	-3.18	-2.84
夫婦と子（未成年）4人	-0.85	-1.04	-0.66
夫婦のみ（高齢）	0.55	0.37	0.72

差異の発生要因に関する分析

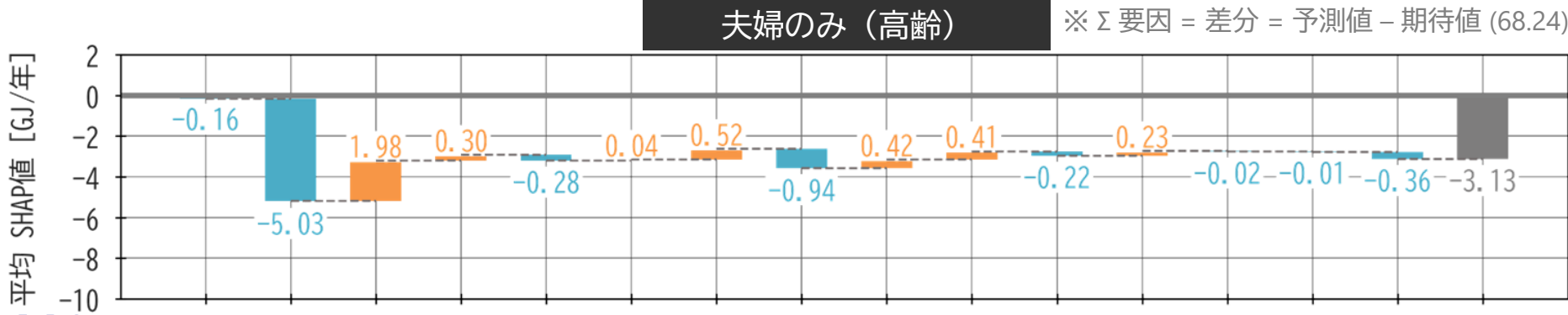
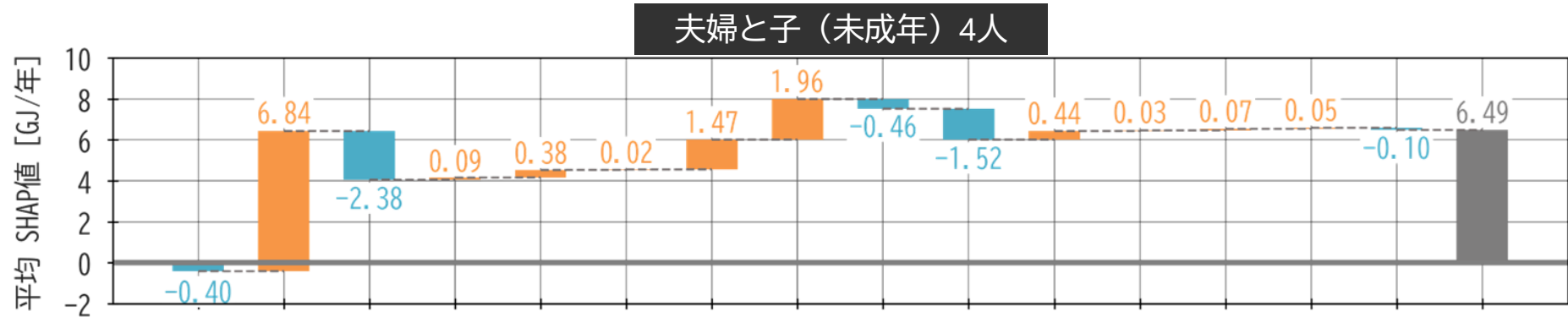
■ 分析結果：要因分解

SHAP値集計：要因ごとに合計+
世帯群ごとに平均
分析内容：予測値の構成要素

- 世帯群によって支配的な要因が異なる。

夫婦と子（未成年）4人： **低下↓** 住宅建物・家電類 **上昇↑** 世帯・給湯・調理

夫婦のみ（高齢）：**低下↓** 世帯 **上昇↑** 住宅建物・給湯・家電類



地域 世帯 住宅建物 暖房 冷房 照明 給湯 調理 テレビ 冷蔵庫 洗濯乾燥 トイレ便座 ペット用品 太陽光発電 省エネ意識 差分

家電類



まとめ

- 環境省が実施する「家庭部門のCO2排出実態統計調査」の調査票情報に基づき、住宅エネルギー消費について次の分析を行った。

統計分析

世帯間における差異に関する分析

- 世帯を世帯類型・居住人数に基づいて分類。
- 年間一次エネルギー消費量および用途別構成比率の群間差について、地域の区分ごとに統計的有意性を検証。
- 地域の区分・世帯群ごとに、用途別構成比率の代表値を推定。

機械学習

差異の発生要因に関する分析

- 機械学習により、予測モデルを作成。
- モデルによる予測結果をSHAPにより解釈し、世帯により差異が生じる要因を把握。

- 一連の分析により、住宅エネルギー消費量に影響を与える要因を特定し、その寄与度を**定量的**に示すことができた。



本研究の成果は、住宅エネルギー消費の実態に対する理解を深めるとともに、さらなる性能向上に向けた対策検討において新たな視点を持つための一助となる。