

全国家計構造調査と 全国単身世帯収支実態調査の 統合集計について2

—都道府県別の2014年全国消費実態調査（単身世帯）
と全国単身世帯収支実態調査の合成—

慶應義塾大学 経済学部/産業研究所

（兼）理研AIPセンター 星野崇宏

慶應義塾大学大学院経済学研究科/産業研究所 二荒麟

研究の概要

- 目的

令和元年度全国家計構造調査の集計に向けて、平成26年度のサンプルサイズは大きいが単身若年層の少ない全国消費実態調査（以下、全消）と、若年層が多い全国単身世帯収支実態調査（以下、全単）のデータを融合¹。

- 注意点 1：両調査の対象者には集団の違い
- 注意点 2：実際の利用時には都道府県別に融合したいという集計スケジュール上のご要望がある
- 手法

前回8月報告時は集団の違いがある可能性があることから強く無視できる割り当てを仮定した、傾向スコアを用いた重みづけを実施

⇒全国サンプルでは精度の向上が見られたが、傾向スコアを用いる際にはサンプルサイズの確保が必要であり、都道府県別に行うことは問題か？

⇒今回は回帰的な手法(回帰代入)も試行

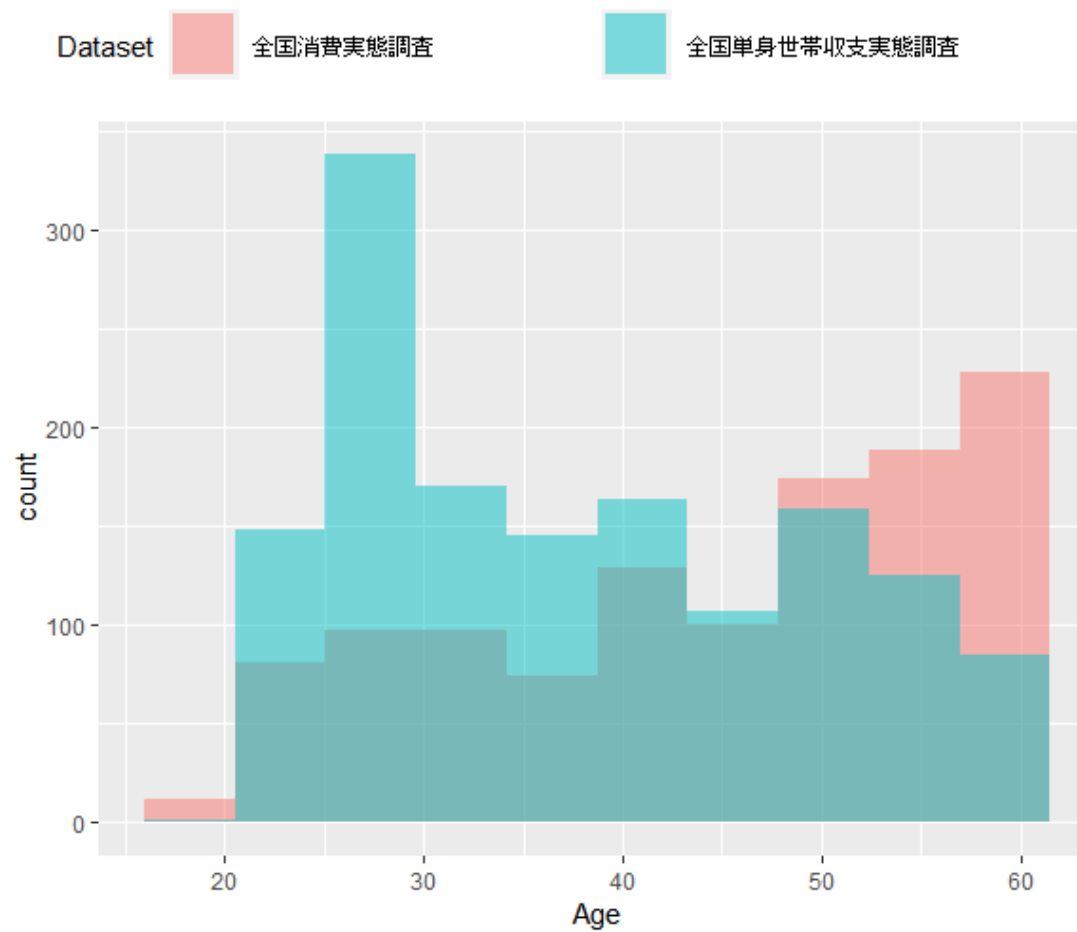
1. ただし、年齢は60歳以下に限定。

地方, 男女, 年齢階級別調査世帯数

		計	30歳未満	30～39	40～49	50～59	60歳以上
男女計	全国	2,000	598	390	290	322	400
	北海道地方	103	28	20	16	18	21
	東北地方	109	32	17	13	23	24
	関東地方	845	257	189	138	120	141
	北陸地方	60	23	8	5	10	14
	東海地方	195	60	40	30	29	36
	近畿地方	324	92	58	48	51	75
	中国地方	101	34	17	12	14	24
	四国地方	54	14	8	4	12	16
	九州地方	191	54	30	21	40	46
	沖縄地方	18	4	3	3	5	3
男	全国	1,144	351	254	198	209	132
	北海道地方	54	16	12	10	10	6
	東北地方	58	17	10	9	15	7
	関東地方	509	156	124	95	81	53
	北陸地方	34	14	5	4	6	5
	東海地方	121	37	30	22	20	12
	近畿地方	178	52	36	33	32	25
	中国地方	55	19	11	8	10	7
	四国地方	28	8	5	3	8	4
	九州地方	97	31	19	12	23	12
	沖縄地方	10	1	2	2	4	1
女	全国	856	247	136	92	113	268
	北海道地方	49	12	8	6	8	15
	東北地方	51	15	7	4	8	17
	関東地方	336	101	65	43	39	88
	北陸地方	26	9	3	1	4	9
	東海地方	74	23	10	8	9	24
	近畿地方	146	40	22	15	19	50
	中国地方	46	15	6	4	4	17
	四国地方	26	6	3	1	4	12
	九州地方	94	23	11	9	17	34 ³
	沖縄地方	8	3	1	1	1	2

総務省統計局が指示する
地域別調査世帯配分数

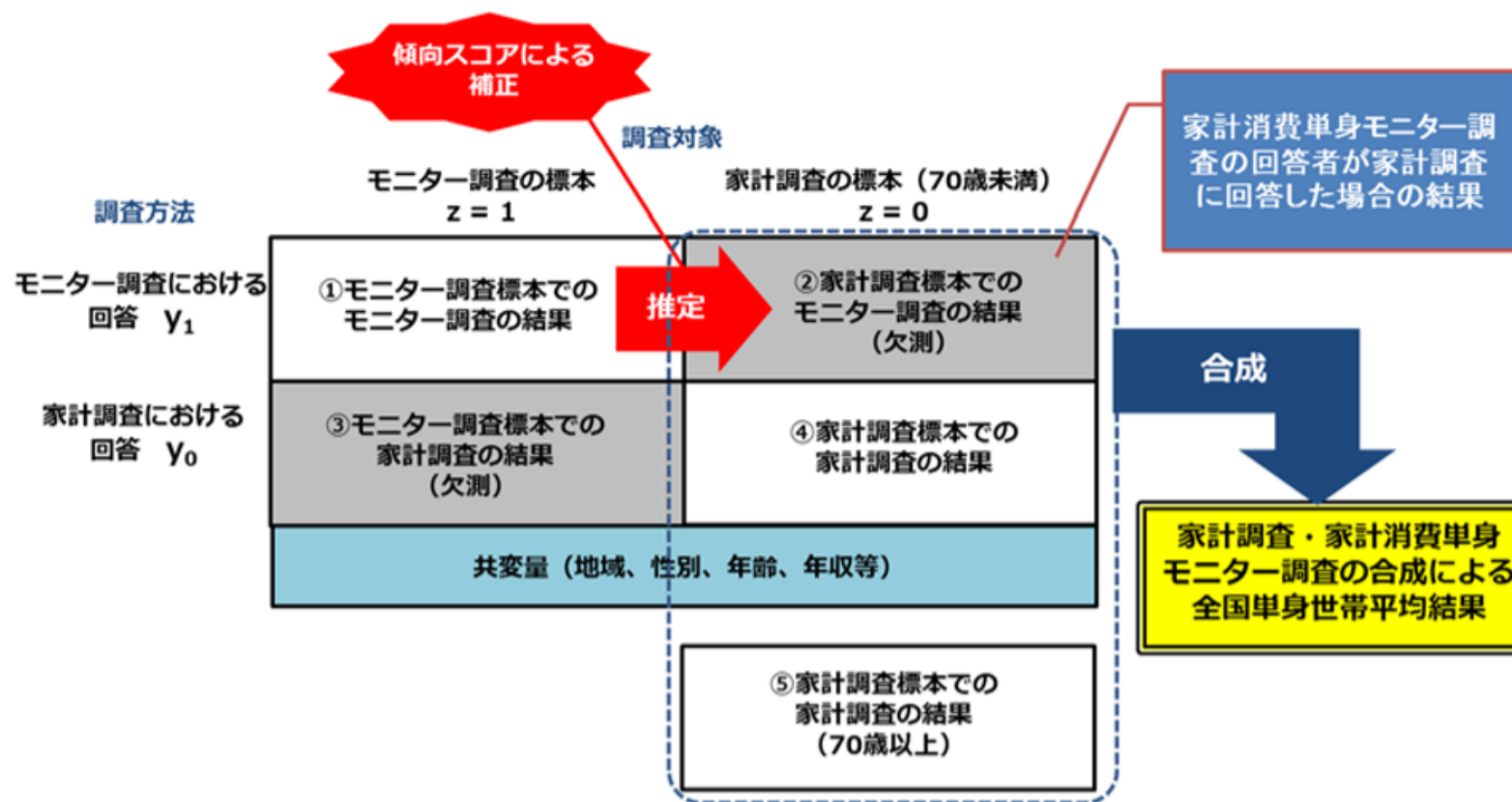
全国での年齢の分布



世帯消費動向指数（CTIマイクロ）のモニター融合

<https://www.stat.go.jp/data/cti/pdf/micro20180309.pdf>

2400人程度の単身モニター対象者の偏りを考慮



手法のイメージ

“選択バイアス” = 回答集団の違い

“調査・データ取得モードの違い” = 取り方の違い

⇒両者の違いが交絡しているので

分離して議論したい

回答集団の違い

今回の目的は
全単調査を用いて
こちらのサンプル
を増やすこと

調査モードの違い

全単調査の
回答結果 y_1

全消調査の
回答結果 y_0

補助変数・
共変量

全単調査の標本
($z = 1$) $N = 1442$



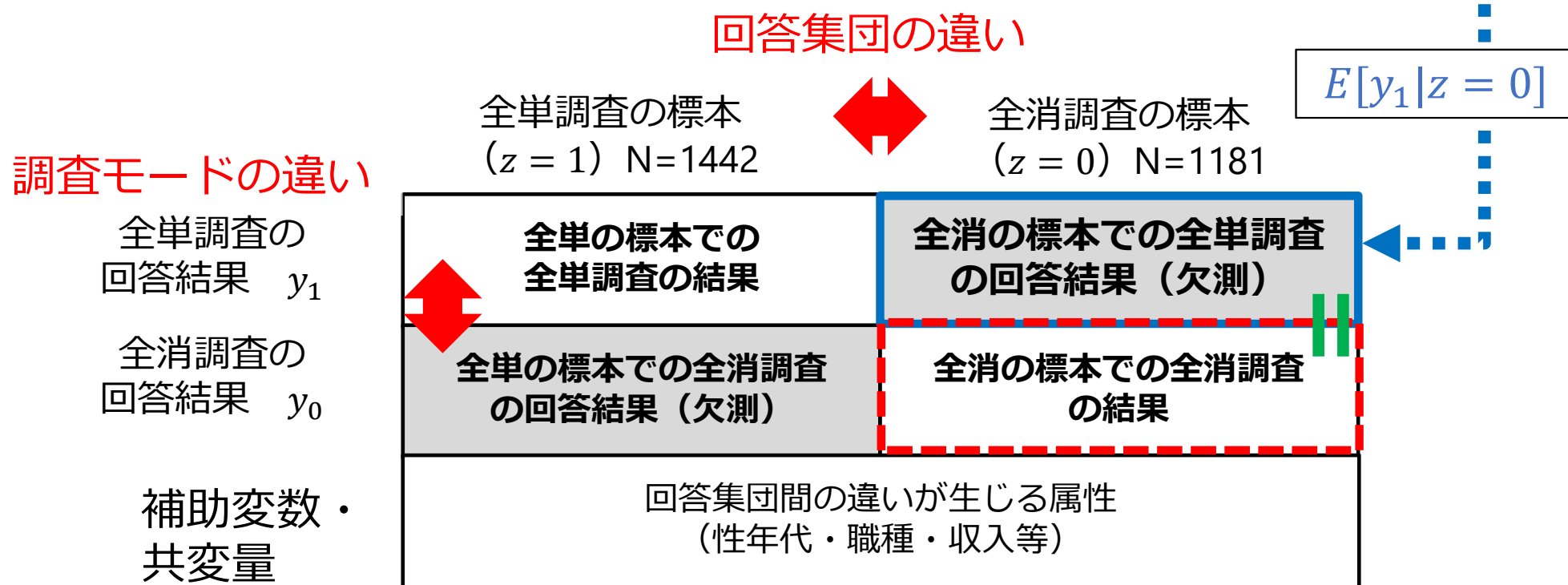
全消調査の標本
($z = 0$) $N = 1181$

全単の標本での 全単調査の結果	全消の標本での全単調査 の回答結果 (欠測)
全単の標本での全消調査 の回答結果 (欠測)	全消の標本での全消調査 の結果
回答集団間の違いが生じる属性 (性年代・職種・収入等)	

手法のイメージ

“調査・データ取得モードの違い”
= 取り方の違い（今回は項目の違い）
がないと仮定すると

もしも年齢の偏りが少ない全単の標本が、全消の標本だった場合の結果
→ 傾向スコア $p(z = 1|x)$
による重みづけ or 回帰的手法



手法のイメージ

“調査・データ取得モードの違い”

= 取り方の違い (今回は項目の違い)

がないと仮定すると

回答集団の違い

全消調査の標本と
して統合 ($z = 0$)

全単調査の標本
($z = 1$) $N=1442$



全消調査の標本
($z = 0$) $N=1181$

全単調査の
回答結果 y_1

全単の標本での
全単調査の結果

全消の標本での全単調査
の回答結果 (欠測)

全消調査の
回答結果 y_0

全単の標本での全消調査
の回答結果 (欠測)

全消の標本での全消調査
の結果

補助変数・
共変量

回答集団間の違いが生じる属性
(性年代・職種・収入等)

【利点】

- ・ サンプルサイズが増える⇒
標本誤差が小さくなる
- ・ 地域別や都市圏別などにして
集計可能

今回比較した手法について

1) 回帰代入の考え方に基づく方法

* 本来はカーネル回帰や部分線形回帰・一般化加法モデルなどもあるがサンプルサイズが小さいので利用できず線形モデルを使用

2) 傾向スコアを用いた重み付け法

利用しなかった手法

- ・ マッチング法(距離による方法)

最近傍・k近傍・予測平均マッチングなど

⇒ 都道府県別では全単・全消のサンプルサイズが少なすぎるため

手法の説明

- Goal: $E[y_1|z = 0]$ の推定

(もしも年齢の偏りが少ない全単の標本が, 全消の標本だった場合の結果)

【方法 1】 傾向スコアによる重み付け (前回8月報告時の方法)

- $\hat{\mu}_1 = \frac{\sum_{i=1}^N \frac{z_i(1-e_i)}{e_i} y_{1i}}{\sum_{i=1}^N \frac{z_i(1-e_i)}{e_i}}$ によって推定.

*IPW(Inverse Probability Weighting)

ただし, e_i は傾向スコア $p(z = 1|x_i)$ の推定値.

【方法 2】 回帰代入の利用

$E[y_1|z = 1, x] = g(x)$ とするとき

- $\hat{\mu}_2 = \frac{1}{N-n} \sum_{i=1}^N (1 - z_i) g(x_i)$ によって推定 (nはz=1のサンプル)

手法のイメージ

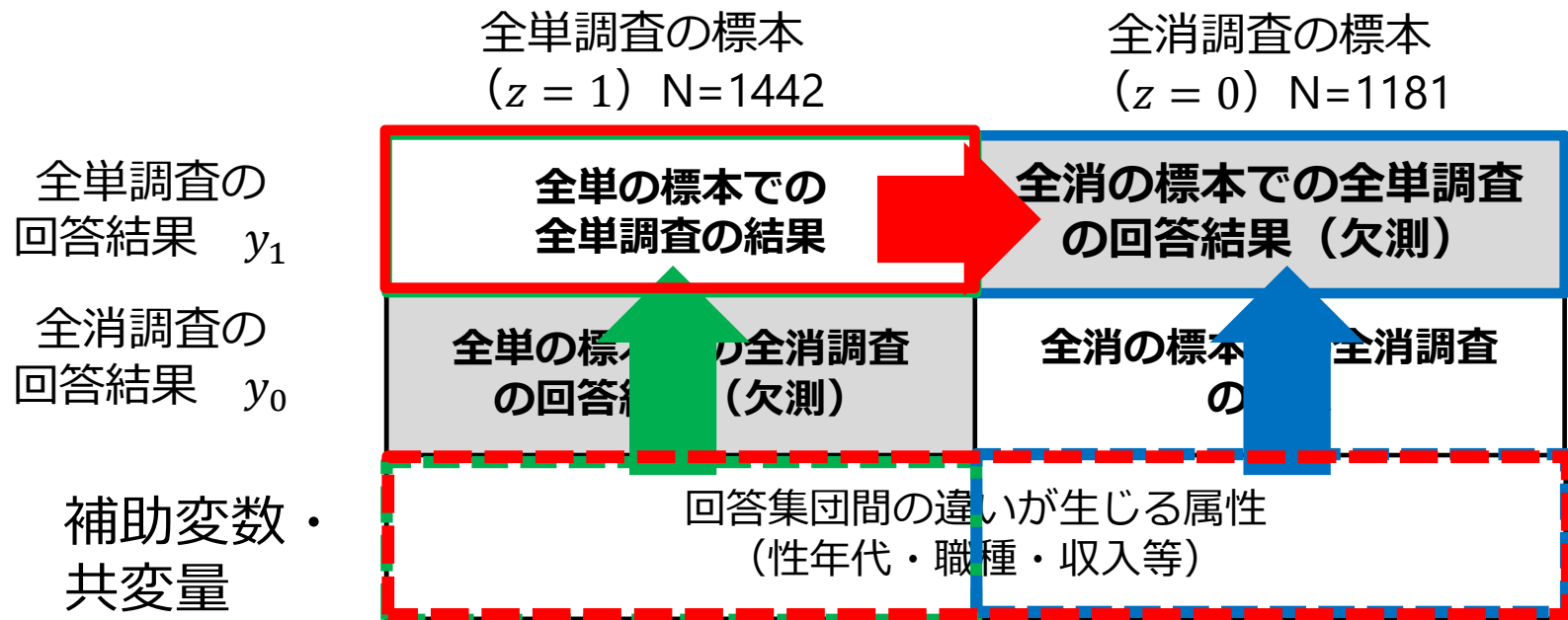
回帰代入*：標本調査法での二相抽出の発想 但し回帰モデル利用
全消の共変量の線形結合(係数は全単のデータ利用)

傾向スコア：標本調査法でのHorvitz-Thompson推定の類推

全単の結果の重み付け

* 厳密には代入法ではない。ここでは公表した集計値や統計量とその標準誤差を計算できればよい

c.f.) 代入法：欠測したデータの値一つ一つの予測値を代入する方法
* 従って多重代入法は不要



両手法の仮定として

ランダムな欠測の仮定(Missing at random)

Rubin(1976) Little and Rubin(2019)

欠測するかしないかが観測された共変量に依存

* 欠測の処理についての標準的な仮定

ここでは

「全国消費実態調査への対象者になるか」「全国単身世帯収支
実態調査の対象者になるか」が利用した共変量に依存する

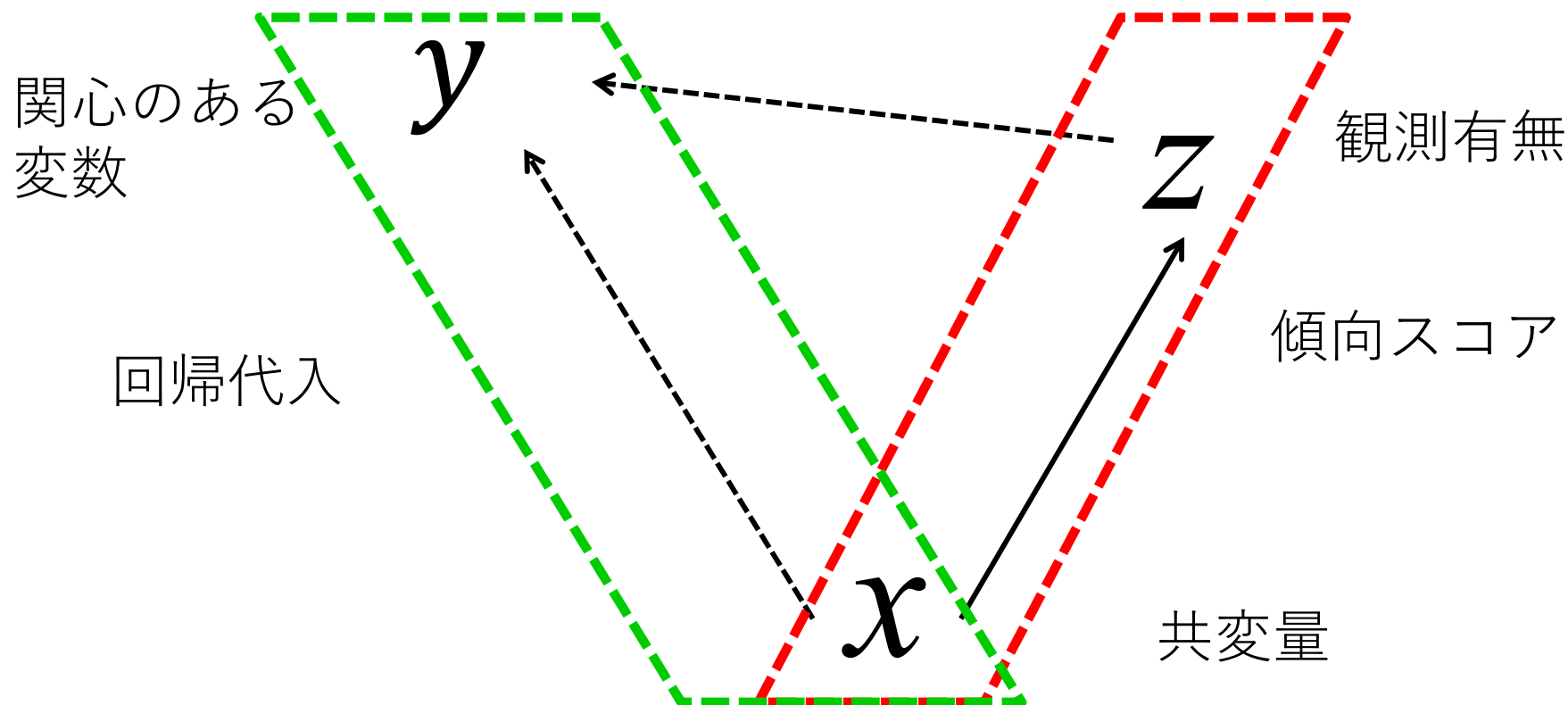
という仮定

このとき

$$E[y_1|z = 1, x] = E[y_1|z = 0, x] = g(x)$$

* 結果として推定量は標本調査法でのDouble samplingの発想を踏まえた手法になる(無限母集団か有限母集団かの違い, Matloff, 1981; Cheng, 1994)

問題の構造



欠測値のある状況での正しい推定のためには「関心のある変数 Y と共変量 X の回帰モデル」または「観測有無 Z と共変量 X の(離散)回帰モデル」どちらかを正しく推定できればよい

方法 2 の推定量とその分散

回帰代入による平均値の推定量は具体的には

$$\frac{1}{N-n} \sum_{i=1}^N (1 - z_i) x_i^t (X^t X)^{-1} X^t y = \frac{1}{N-n} \sum_{i=1}^N (1 - z_i) x_i^t \hat{\beta}$$

但し $\hat{\beta}$ は $z=1$ で共変量による従属変数の回帰分析をした最小二乗推定量
またその分散は以下の式による計算

$$\frac{s^2}{(N-n)^2} \mathbf{1}^t X_0 (X^t X)^{-1} X_0^t \mathbf{1}$$

但し y は $z=1$ のみでの回答者の従属変数を集めたもの

同様に X は $z=1$ のみでの回答者の共変量ベクトルを集めたもの

n は $z=1$ のサンプルサイズ, s^2 は $z=1$ での残差分散の不偏推定量

X_0 は $z=0$ での回答者の共変量ベクトルを集めたもの

$\mathbf{1}$ は $N-n$ 次元で要素がすべて1のベクトル

合成値の考え方

調査モードがないとするので $\theta_0 = E[y_0|z=0] = E[y_1|z=0]$ には二つの推定量 \bar{y}_0 と $\hat{E}[y_1|z=0]$ が存在。これらに重み w をつけて

$$\hat{\theta}_0 = w\bar{y}_0 + (1-w)\hat{E}[y_1|z=0] \text{ とするときの } w \text{ は？}$$

【1:最適な重みにする】 \Rightarrow 今回はこちら (なるべく分散を小さくする)

$V(\hat{\theta}_0)$ を最小にする重みは (両推定量がほぼ無相関のため)

$$w = \frac{V(\hat{E}[y_1|z=0])}{V(\bar{y}_0) + V(\hat{E}[y_1|z=0])}$$

【2:人数比で考える】

全消調査の標本のサンプルサイズ ($z=0$) $N_0 = 1181$ と

全単調査の標本のサンプルサイズ ($z=1$) $N_1 = 1442$ より

$$w = \frac{N_0}{N_0 + N_1}$$

* その後の層別の分析の際の重みの分かりやすさという点では2か？

今回の調整のための変数の候補

消費統計研究会(平成29年度第1回)

資料4 別紙「傾向スコアを用いた家計調査（単身世帯）と家計消費単身モニター調査の合成に関する試算結果」に存在する変数を利用

性別ダミー 年齢（連続） 居住面積（連続）

世帯年収の対数値（連続） 純資産総額（連続） ⇒年間収入で使用せず

消費支出の対数値（連続） ⇒消費支出で使用せず

職業分類区分（13区分⇒5分類に）

住居の所有関係（6区分⇒3分類に） 都市規模（5区分⇒4分類に）

地方区分（10区分⇒5分類に）

上記フルセットの場合ダミー変数を入れて

消費支出では19変数 年間収入では18変数

variable	変数	符号表の名前	値	説明	分析用の区分	変数
gender	性別	Seibetsu	1	男		1
			2	女		0
age	満年齢	Nenrei	△△0~116	0~116歳		
job	職業符号	Shokugyo	△1	常用労務作業者	労務作業者	JOB1
			△2	臨時及び日々雇労務作業者	労務作業者	JOB1
			△3	民間職員	民間職員	基準
			△4	官公職員1	官公職員	JOB2
			△5	官公職員2	官公職員	JOB2
			△6	商人及び職人	民間職員	基準
			△7	個人経営者	経営者	JOB3
			△8	農林漁業従事者	民間職員	基準
			△9	法人経営者	経営者	JOB3
			10	自由業者	経営者	JOB3
			11	その他	経営者	JOB3
			12	無職	無職	JOB4
			13	家族従業者	民間職員	基準
housing	住居の所有関係	JukyoShoyu	1	持ち家	持ち家	HOU1
			2	民営の賃貸住宅	民営の賃貸住宅	基準
			3	都道府県・市区町村営賃貸住宅	公営の賃貸住宅・社宅等	HOU2
			4	都市再生機構・公社等の賃貸住宅	公営の賃貸住宅・社宅等	HOU2
			5	社宅・公務員宿舎(借上げの社宅を含む)	公営の賃貸住宅・社宅等	HOU2
			6	借間	民営の賃貸住宅	基準
space	住居の敷地面積	ShikichiMenseki	△△△0.1~9999.9	住居の敷地面積		
city	都市階級	C_ToshiKaikyuu	1	大都市	大都市	基準
			2	中都市	中都市	CIT1
			3	小都市A	小都市A	CIT2
			4	小都市B	小都市B・町村	CIT3
			5	町村	小都市B・町村	CIT3
region	地方区分	ChihoKubun	01	北海道	北海道・東北	REG1
			02	東北	北海道・東北	REG1
			03	関東	関東	基準
			04	北陸	北陸・東海	REG2
			05	東海	北陸・東海	REG2
			06	近畿	近畿	REG3
			07	中国	中国・四国	REG4
			08	四国	中国・四国	REG4
			09	九州	九州・沖縄	REG5
			10	沖縄	九州・沖縄	REG5

consum	消費支出	Bdy083	△△△△ △△△ 0.000000 ~	金額(円)
income	世帯の年間収入	M_Nenshu	△△△△ △△△△ △1~	金額(万円)
asset	純資産総額	BdyShisan011	- 9999999.9 99999~	金額(千円)
debt	負債現在高	BdyShisan020	△△△△ △△△ 0.000000 ~	金額(千円)
saving	貯蓄計	ChochikuTotal	△△△△ △1~ 999999	金額(万円)
engel	エンゲル係数			consum/income
ln_con	消費支出の対数値			log(consum)
ln_inc	世帯年収の対数値			log(income)
ln_sav	貯蓄計の対数値			log(saving)

前提としての都道府県別のサンプルサイズ

都道府県別に全単と全消を融合させる場合

	北海道	青森県	岩手県	宮城県	秋田県	山形県	福島県	茨城県	栃木県	群馬県	埼玉県	千葉県	東京都	神奈川県	新潟県	富山県
全単	74	14	14	23	7	8	16	26	16	26	71	68	269	135	18	10
全消	44	15	16	24	19	15	28	28	16	17	67	61	60	64	30	11
	石川県	福井県	山梨県	長野県	岐阜県	静岡県	愛知県	三重県	滋賀県	京都府	大阪府	兵庫県	奈良県	和歌山県	鳥取県	島根県
全単	10	4	9	25	14	31	75	18	11	36	97	58	13	7	5	3
全消	18	12	16	27	17	40	58	20	20	11	43	46	8	17	13	17
	岡山県	広島県	山口県	徳島県	香川県	愛媛県	高知県	福岡県	佐賀県	長崎県	熊本県	大分県	宮崎県	鹿児島県	沖縄県	全国
全単	16	36	11	4	9	11	6	57	6	13	11	11	9	18	13	1442
全消	18	23	9	19	16	18	16	30	17	16	22	16	15	20	28	1181

全単のサンプルサイズは島根で3、福井・徳島で4、鳥取で5、高知・佐賀で6

全消のサンプルサイズは奈良で8、山口で9

⇒およそ分析が行えないサイズ 回帰なら最低で説明変数 + 2

* さらに所属のないカテゴリーの存在 (例) 全単山形県持ち家ゼロ人
 ほぼすべての方法は都道府県別では実施できない?

変数群

【フルセット】

消費支出では19変数 年間収入では18変数

* 年間年収の対数値 ・ 純資産総額 ⇒年間収入で使用せず

* 消費支出の対数値 ⇒消費支出で使用せず

【セット1】 カテゴリカル変数を除外

消費支出：性別・年齢・居住面積・世帯年収対数値・純資産総額

年間収入：性別・年齢・居住面積・消費支出

【セット2】 全国データでAICから最良のセット

消費支出：性別・年齢・居住面積・職業分類・住居の所有関係

・ 年間収入の対数値・純資産総額 (都市規模・地域区分削除)

年間収入：性別・年齢・居住面積・職業分類・住居の所有関係・

・ 都市規模・消費支出の対数値 (純資産総額・地域区分削除)

比較した分析方法

消費支出と世帯収入を調整の目的変数として

以下の7タイプの分析を実施

①全国でフルセットで回帰代入/傾向スコア

* 傾向スコアについては8月分析時同様

②全国でセット1で回帰代入/傾向スコア

③全国でセット2で回帰代入/傾向スコア

④地域ごと(北海道・東北・関東・北陸・東海・近畿・中国・四国

・九州沖縄)でフルセットで回帰代入/傾向スコア

⑤地域ごとでセット1で回帰代入/傾向スコア

⑥地域ごとでセット2で回帰代入/傾向スコア

* 地域別推定では都市規模は削除

⑦都道府県別でセット1で回帰代入/傾向スコア

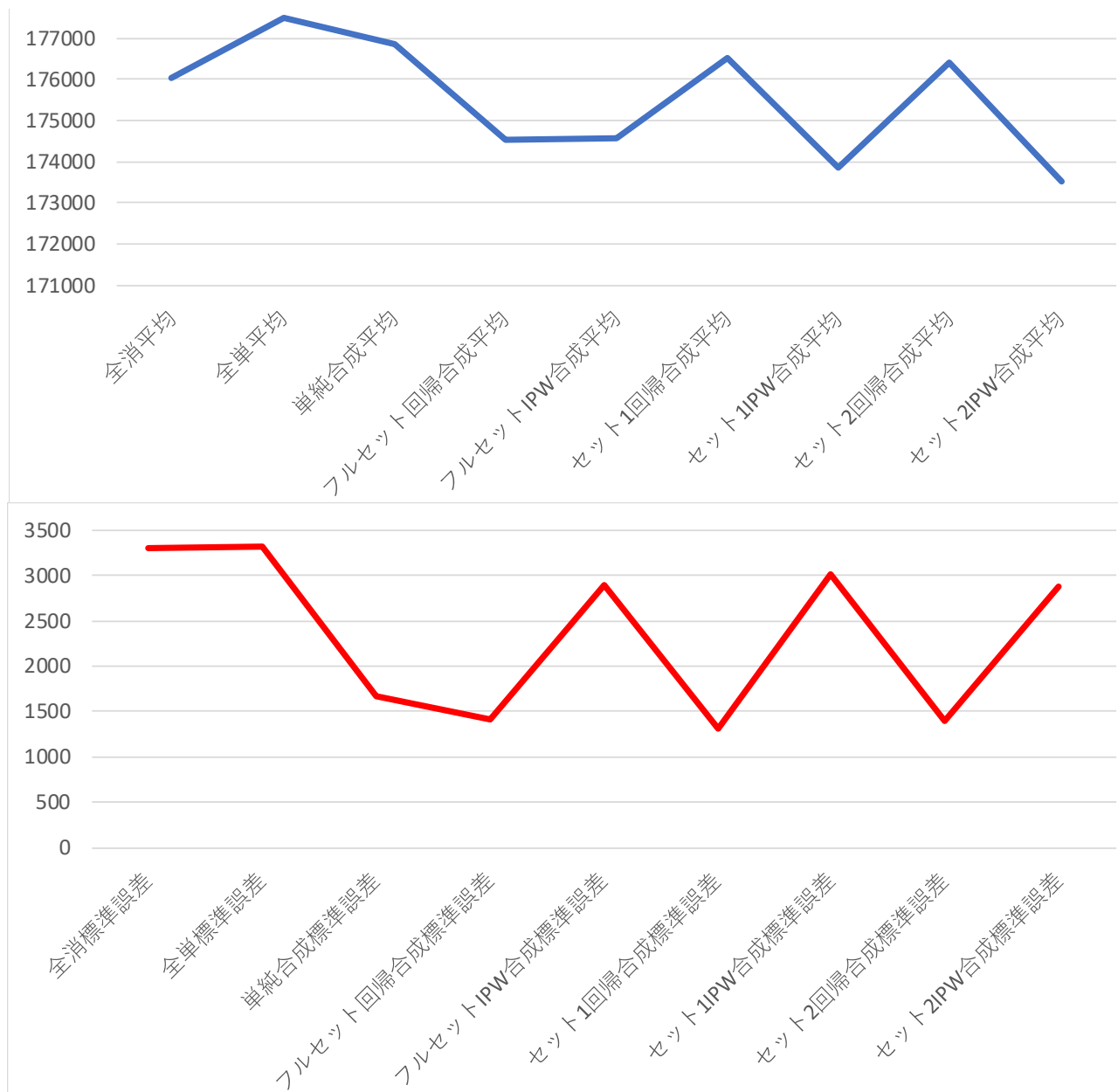
* すでにみたように都道府県別ではフルセットとセット2は実施不可能

消費支出(全国)

①～③

右上：平均値（円）

左上：平均値の標準誤差

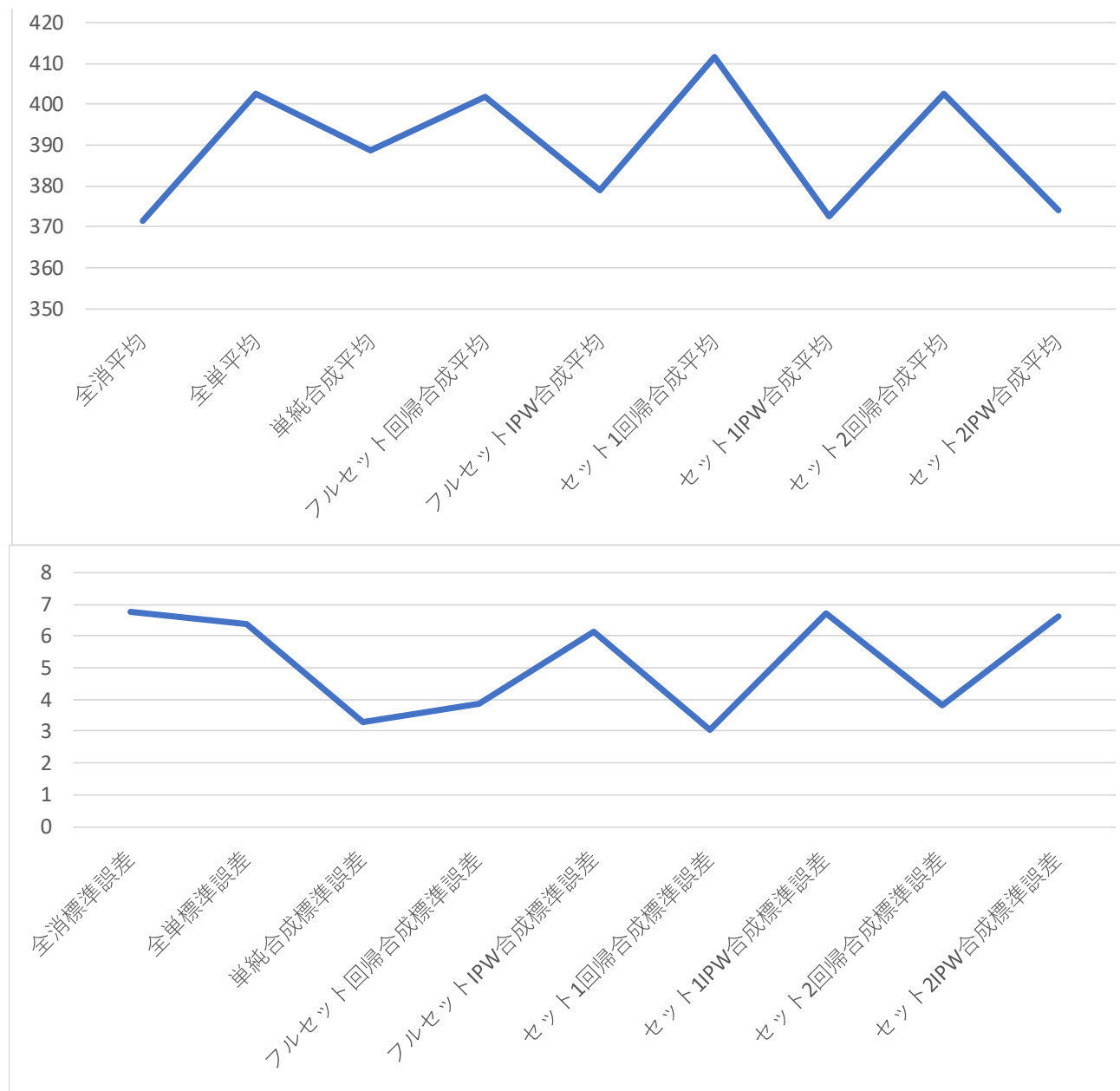


年間収入(全国)

①～③

右上：平均値（円）

左上：平均値の標準誤差

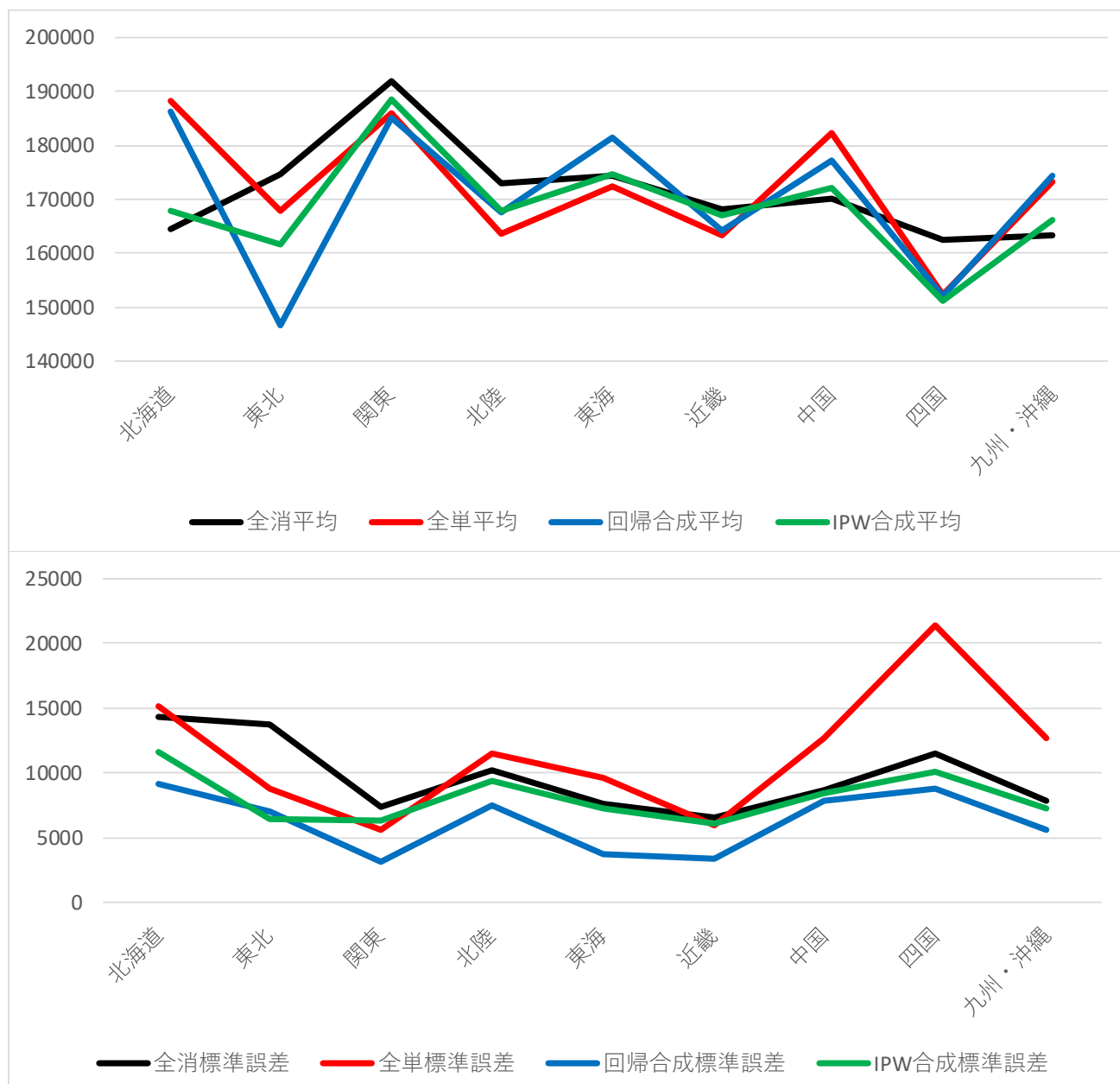


消費支出(地域)

④フルセット

右上：平均値 (円)

左上：平均値の標準誤差

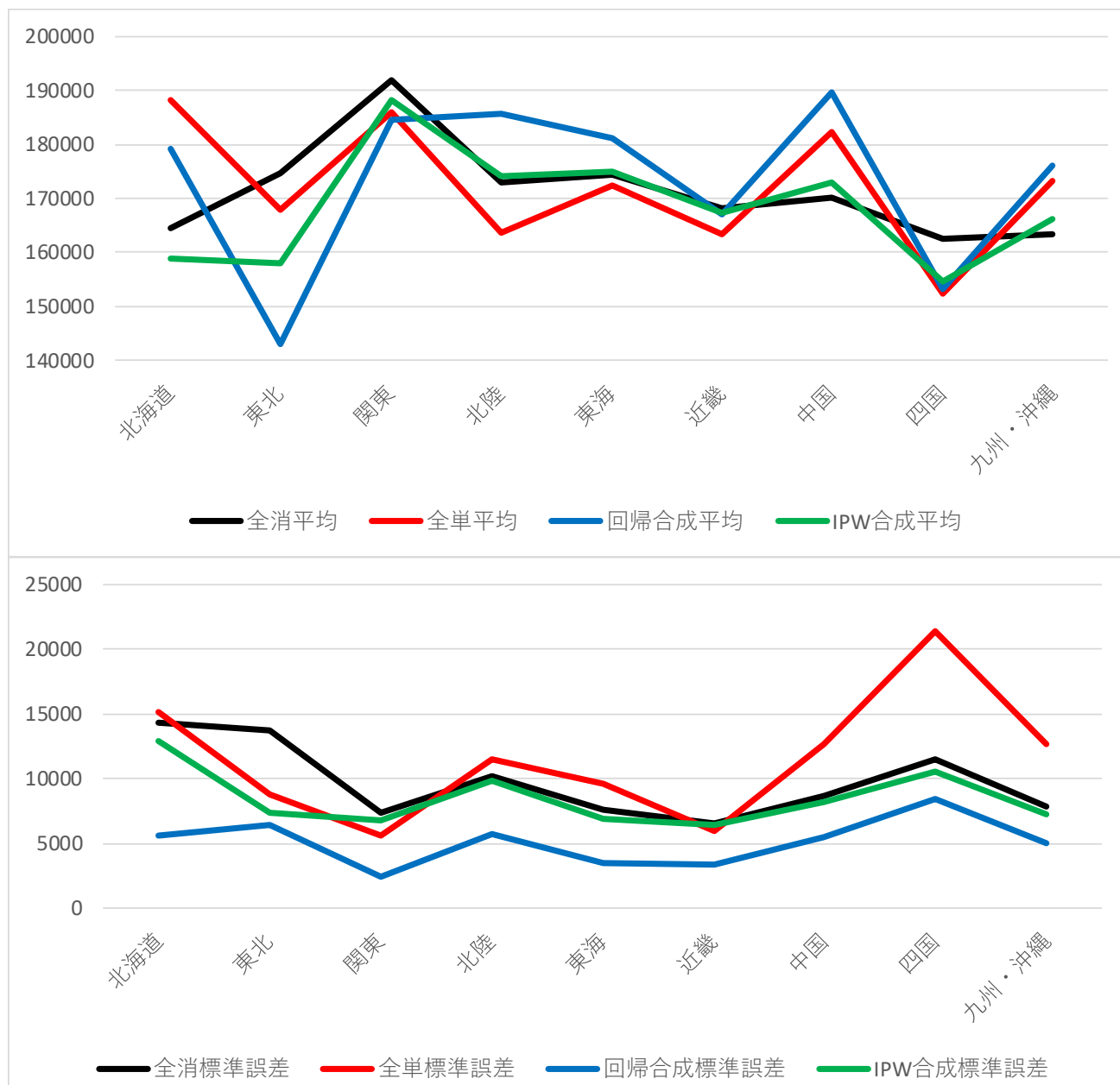


消費支出(地域)

⑤セット1

右上：平均値 (円)

左上：平均値の標準誤差

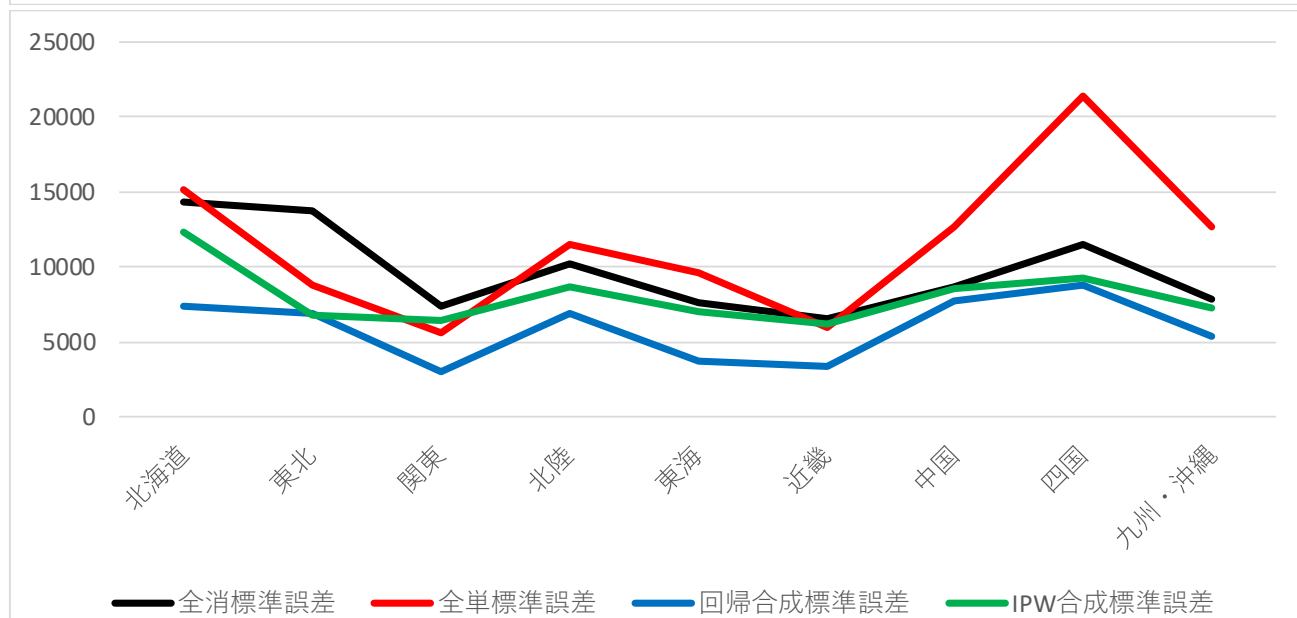
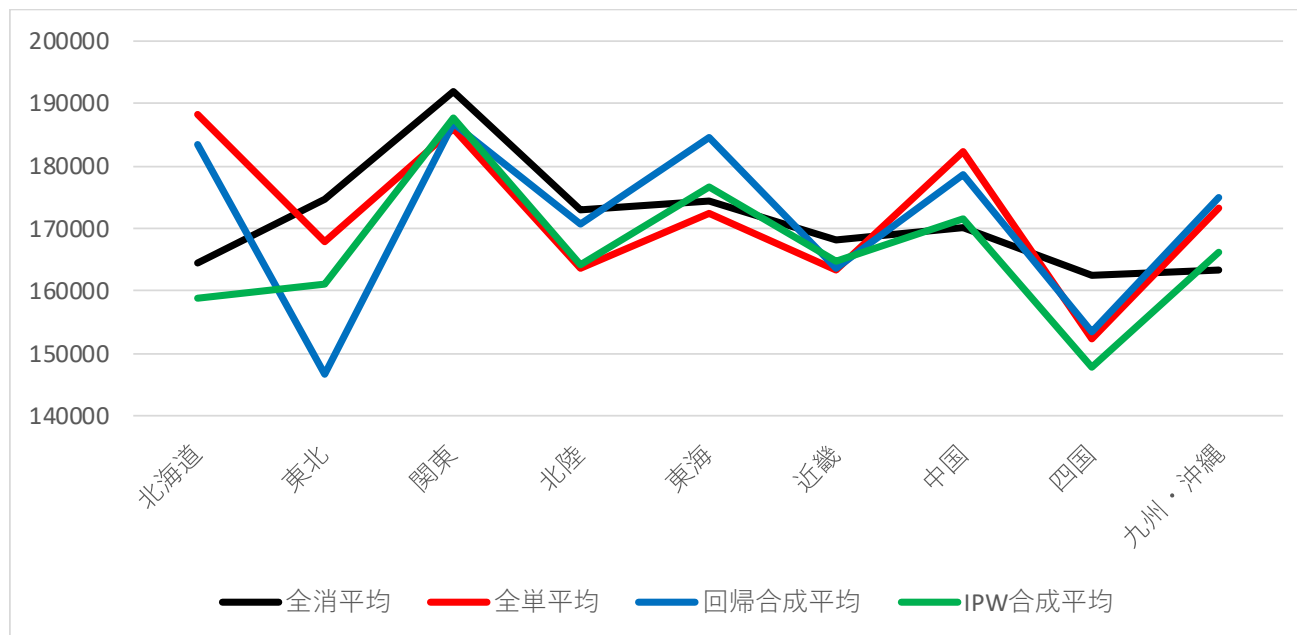


消費支出(地域)

⑥セット2

右上：平均値 (円)

左上：平均値の標準誤差

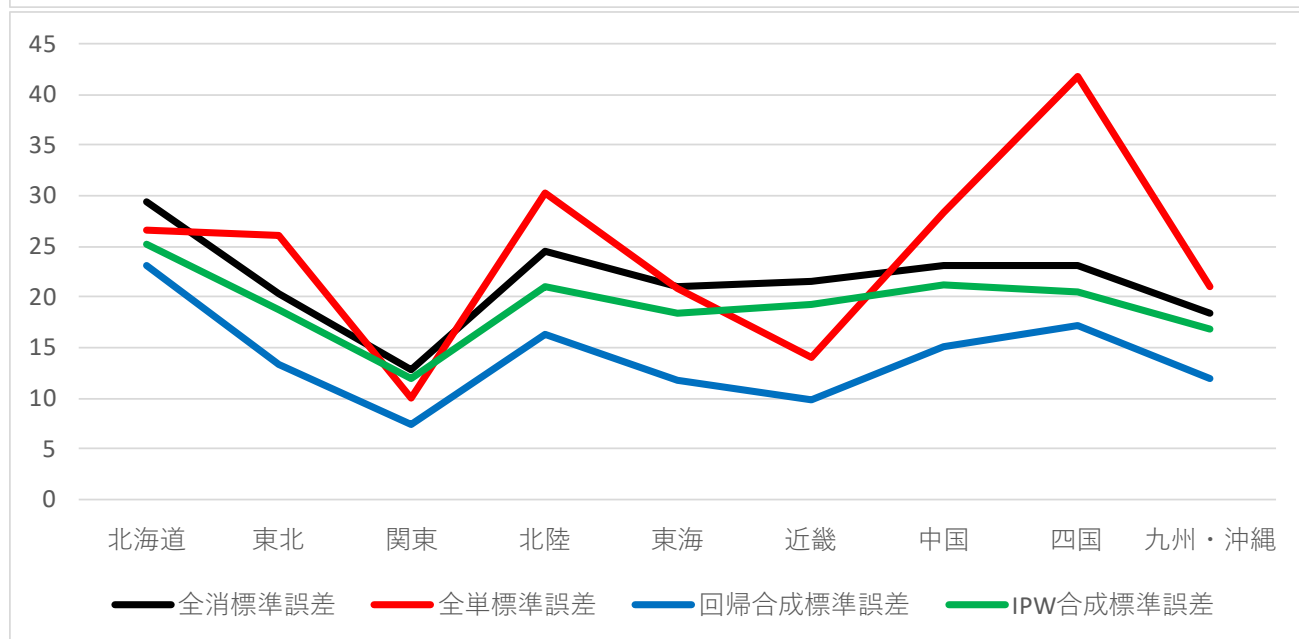
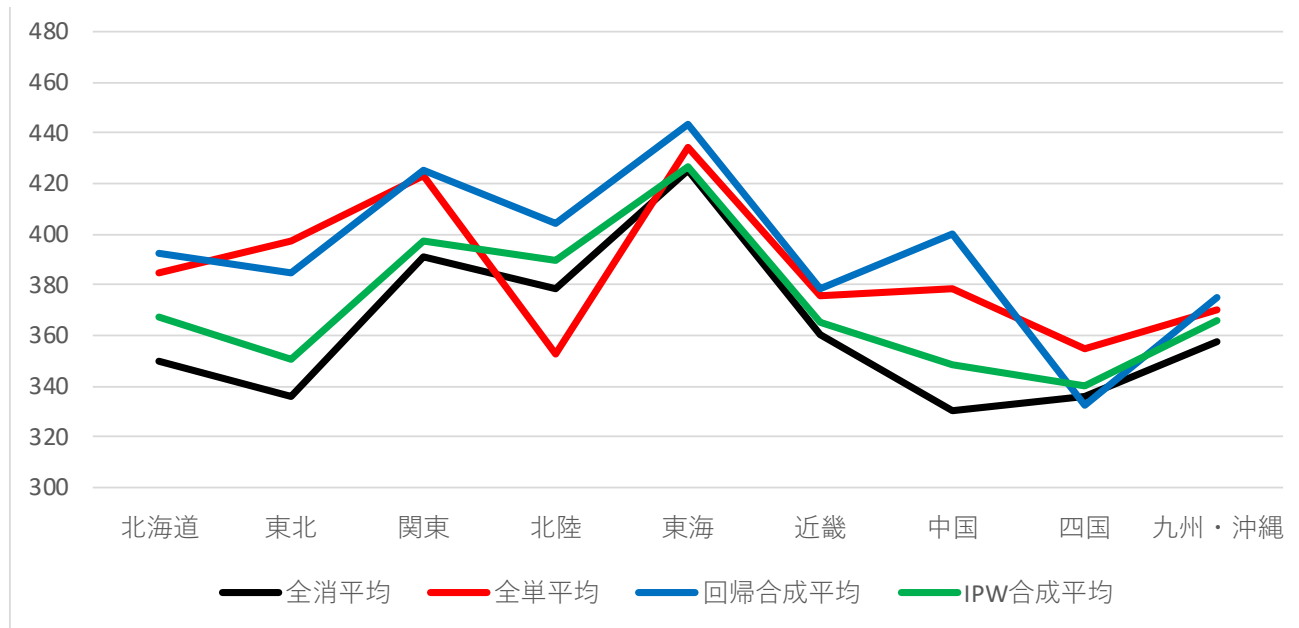


年間収入(地域)

④フルセット

右上：平均値 (円)

左上：平均値の標準誤差

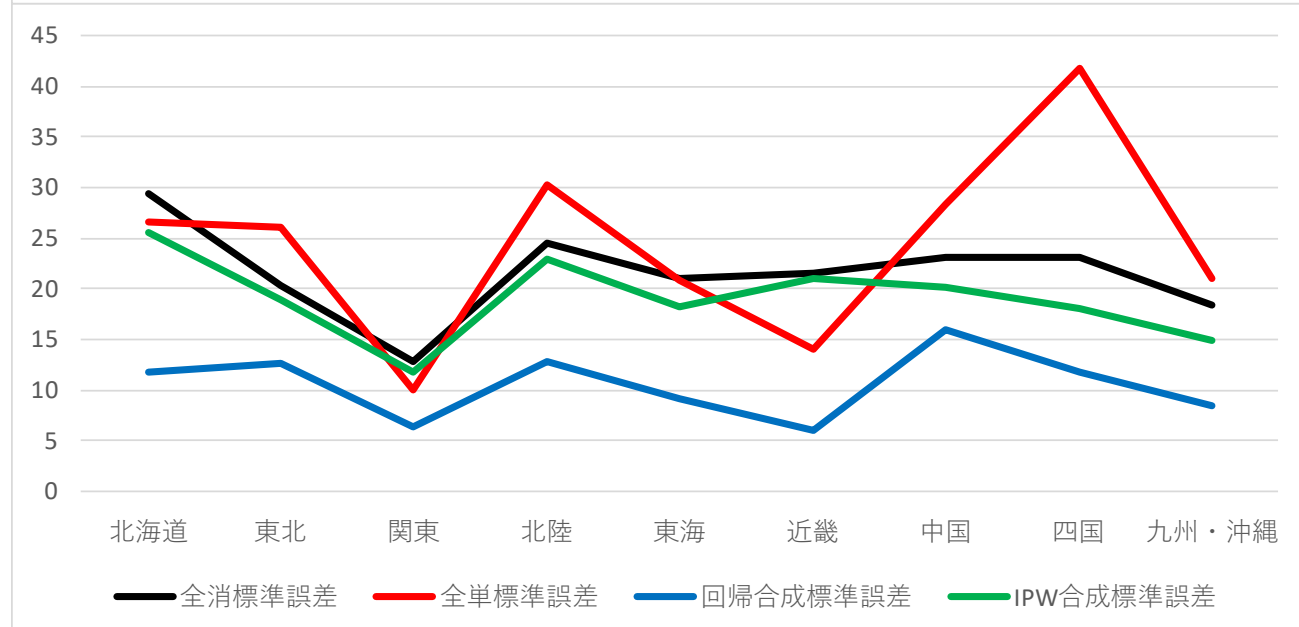
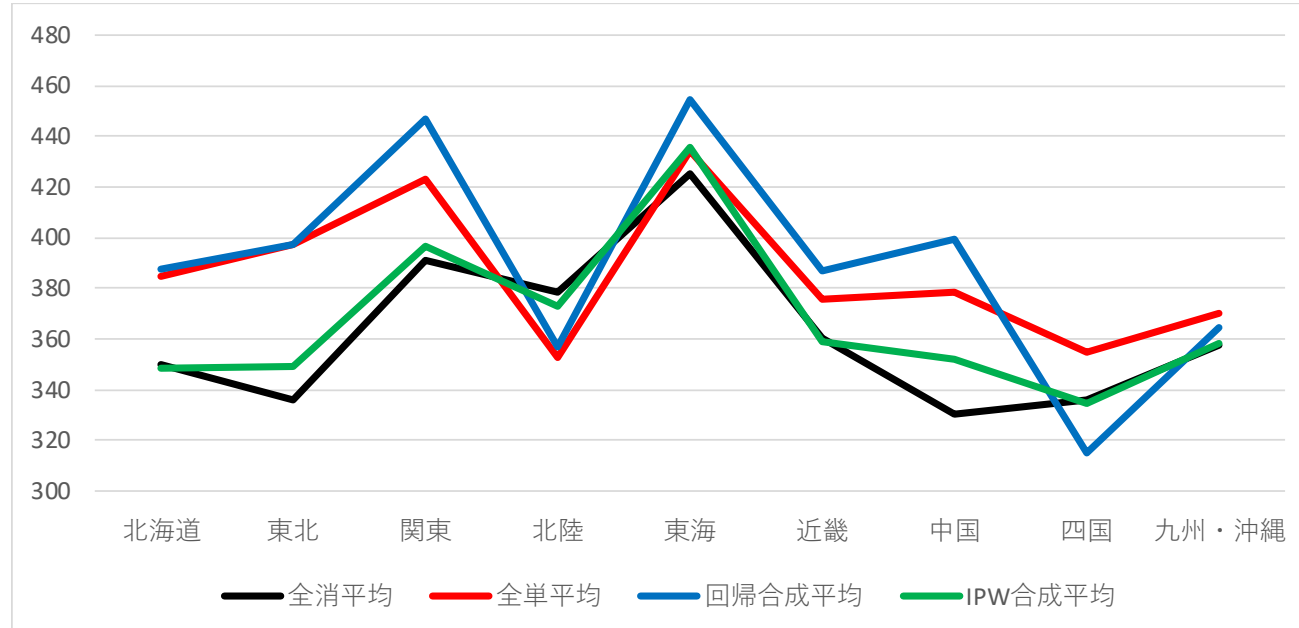


年間収入(地域)

⑤セット1

右上：平均値 (円)

左上：平均値の標準誤差

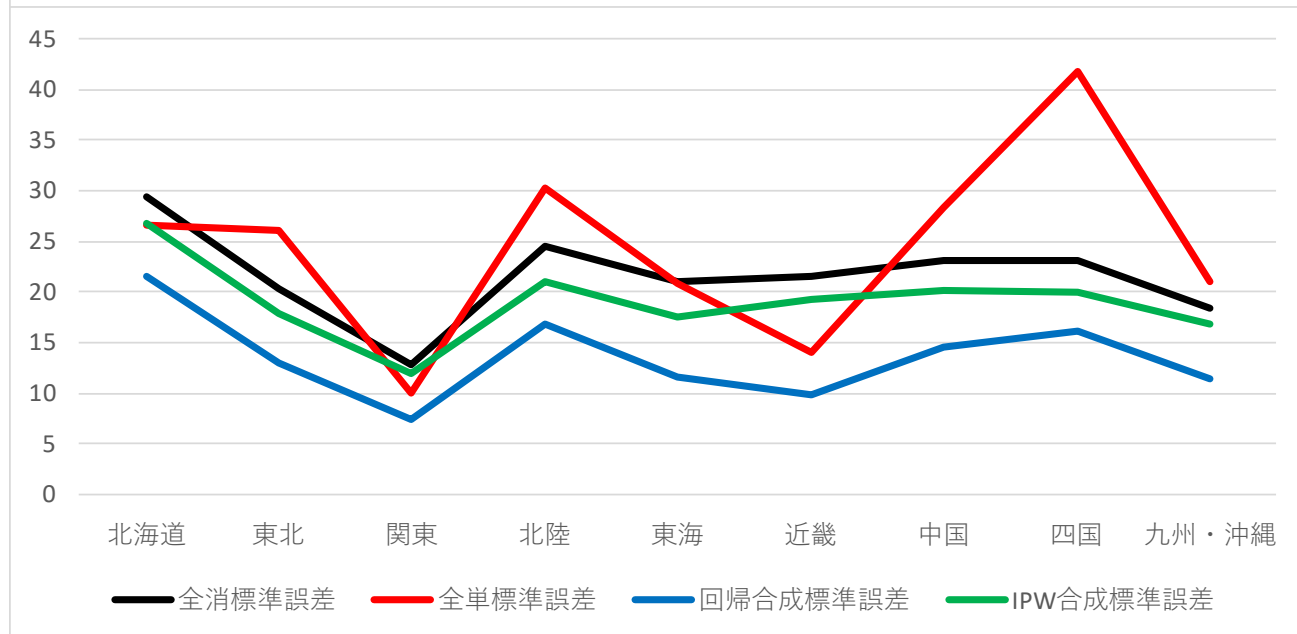
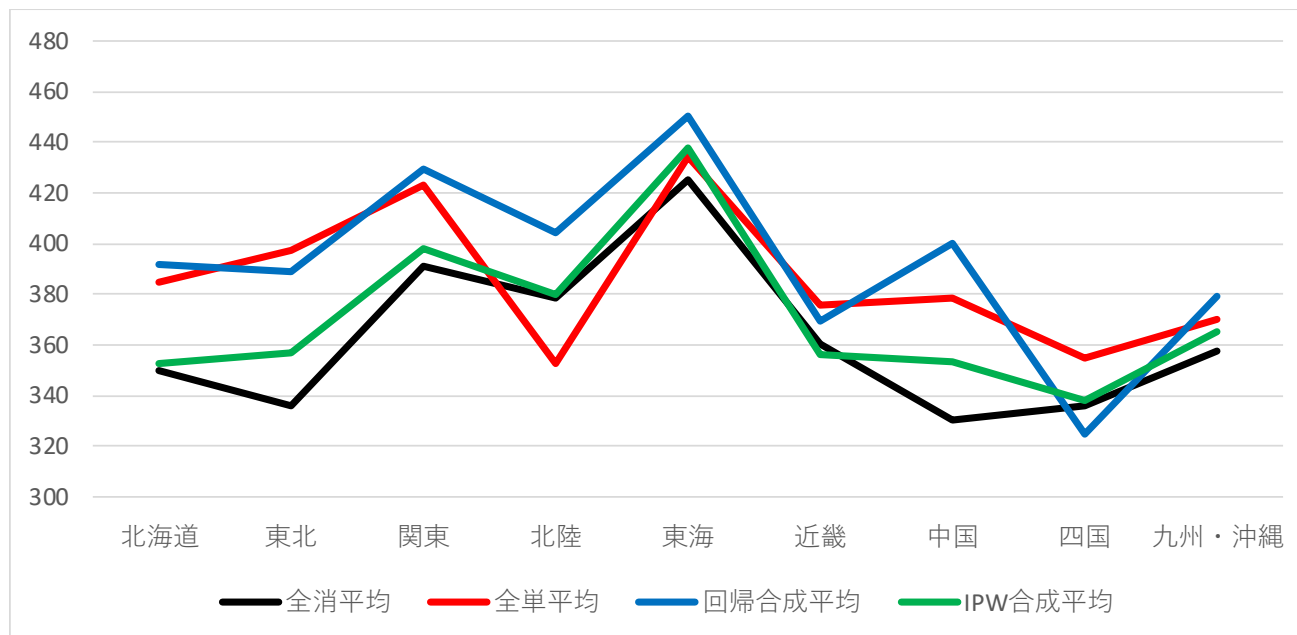


年間収入(地域)

⑥セット2

右上：平均値 (円)

左上：平均値の標準誤差



全国レベル・地域レベルの分析のまとめ

- 全単と全消の2調査の対象者の違いを考慮して調整を行うという点では全単を全消（あるいは今年度の全国家計構造）に取り込むには回帰代入・傾向スコアによるIPWを行う必要がある。
- 取り込んだ後の合成値の標準誤差は当然ながら原則として向上する。
- 2調査の対象者の違いがないという仮定であれば、単純な集計(両者をまとめる)より標準誤差は良くなるわけではない。
 - * データの情報を両者の集団の違いを調整するために利用するため。
- 但し、2調査の対象者が異質であるならば、単純な集計およびその標準誤差は信頼できない。
 - * 両者の差異は全国では統計的に有意な差・地域ごとでも一部有意な差あり

都道府県別の解析について

【回帰代入】

回帰分析を行う全単のサンプルサイズが小さく、変数セットはセット1のみ利用
それでも

消費支出

福井県、島根県、鳥取県、徳島県、高知県、佐賀県で計算できず

年間収入

福井県、島根県、鳥取県、徳島県で計算できず

*参考：全単のサンプルサイズは島根で3、福井・徳島で4、鳥取で5、
高知・佐賀で6

全消のサンプルサイズは奈良で8、山口で9

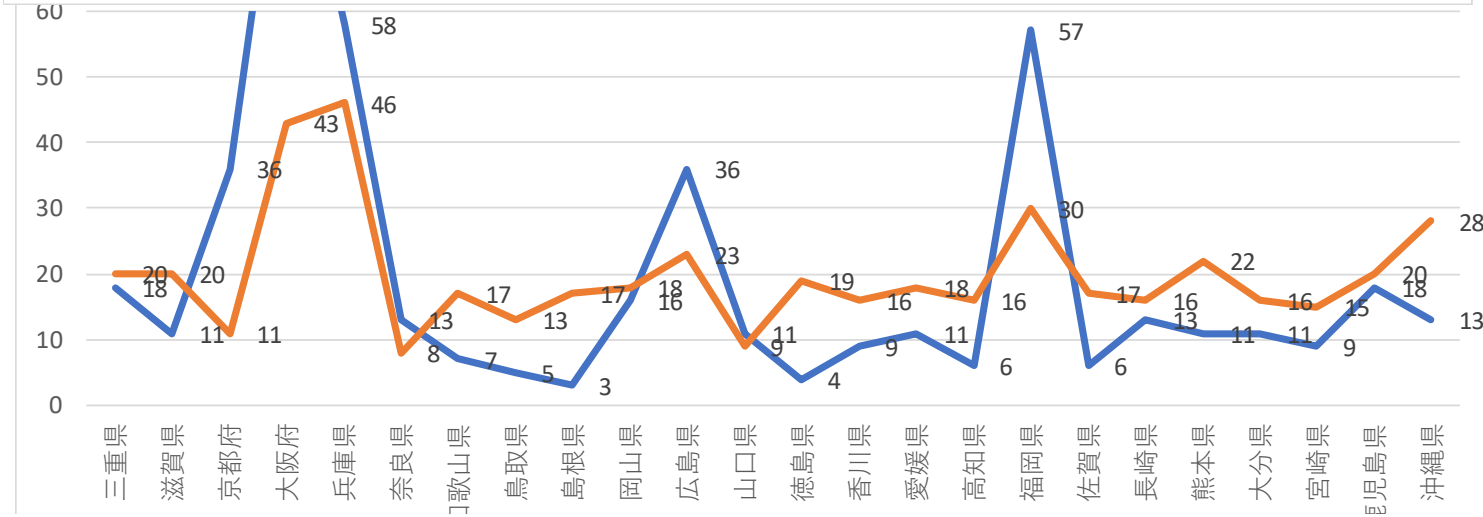
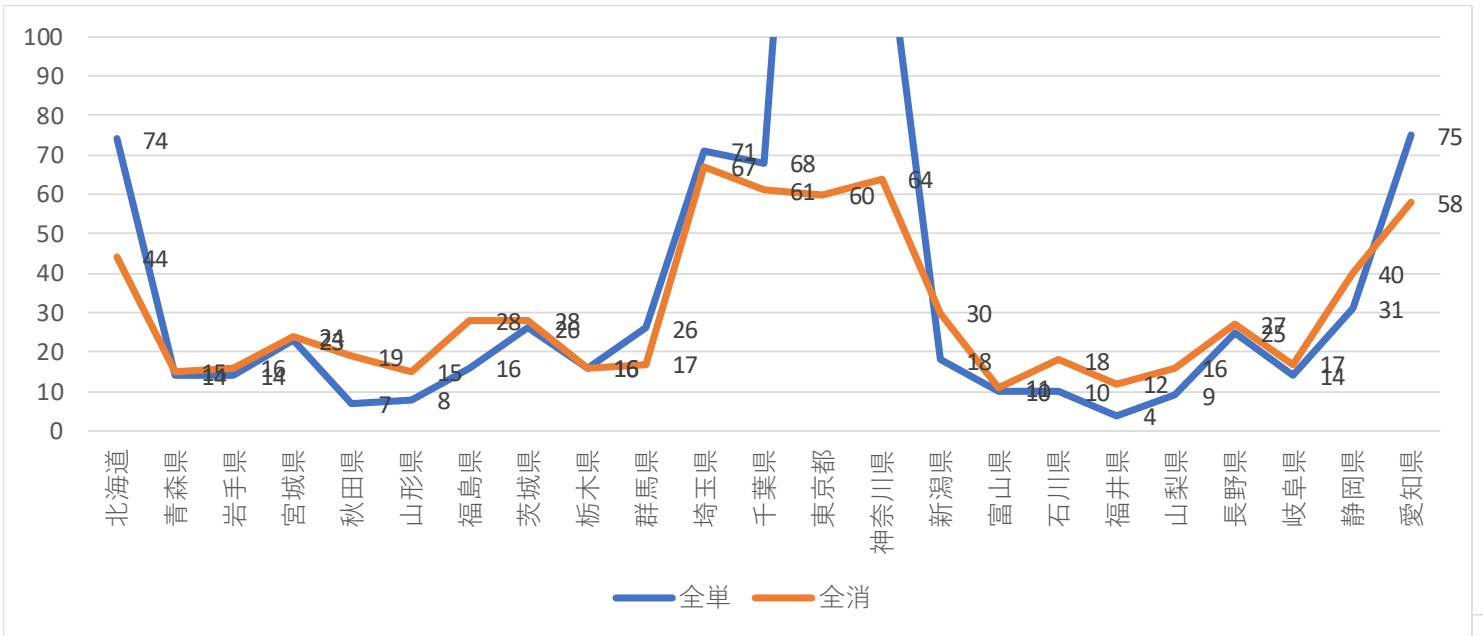
⇒サンプルサイズの順に計算できず

【傾向スコア】

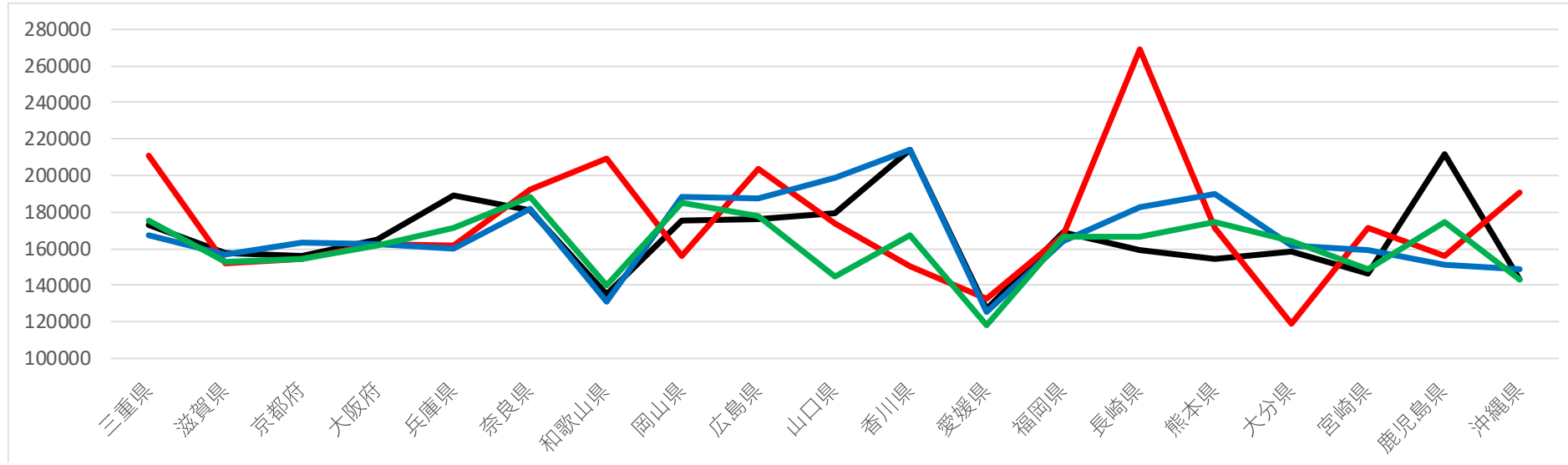
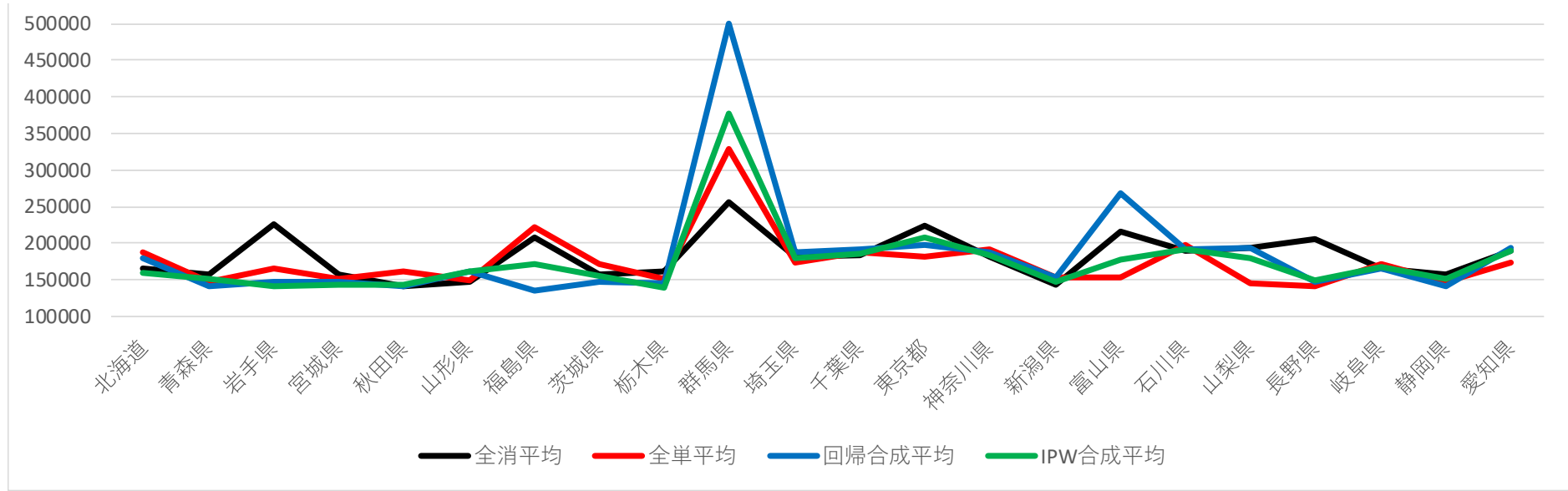
こちらは全単と全消どちらも使うためか一応解析全都道府県で実行可能

サンプルサイズ(都道府県別)

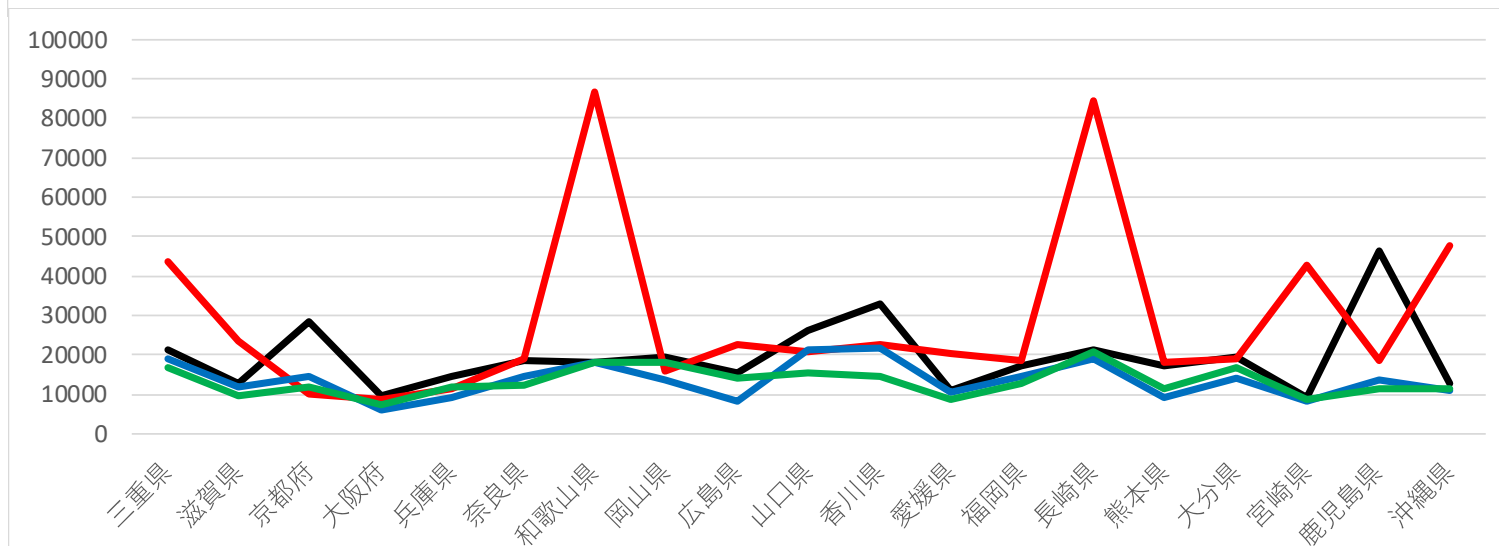
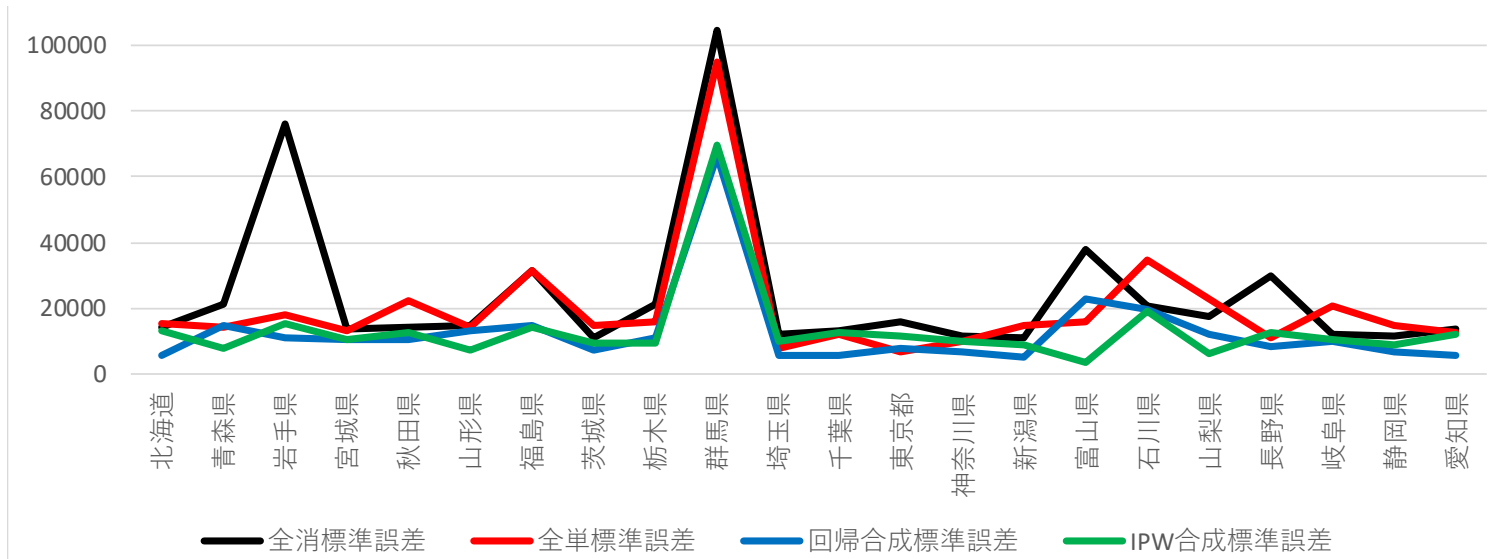
全単
東京は269
大阪は97



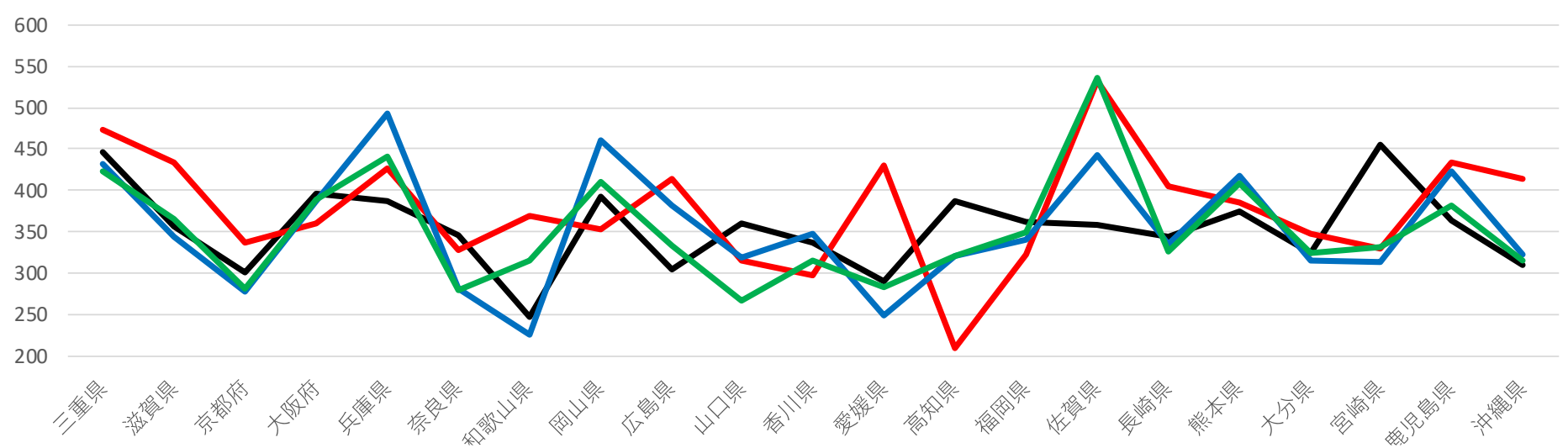
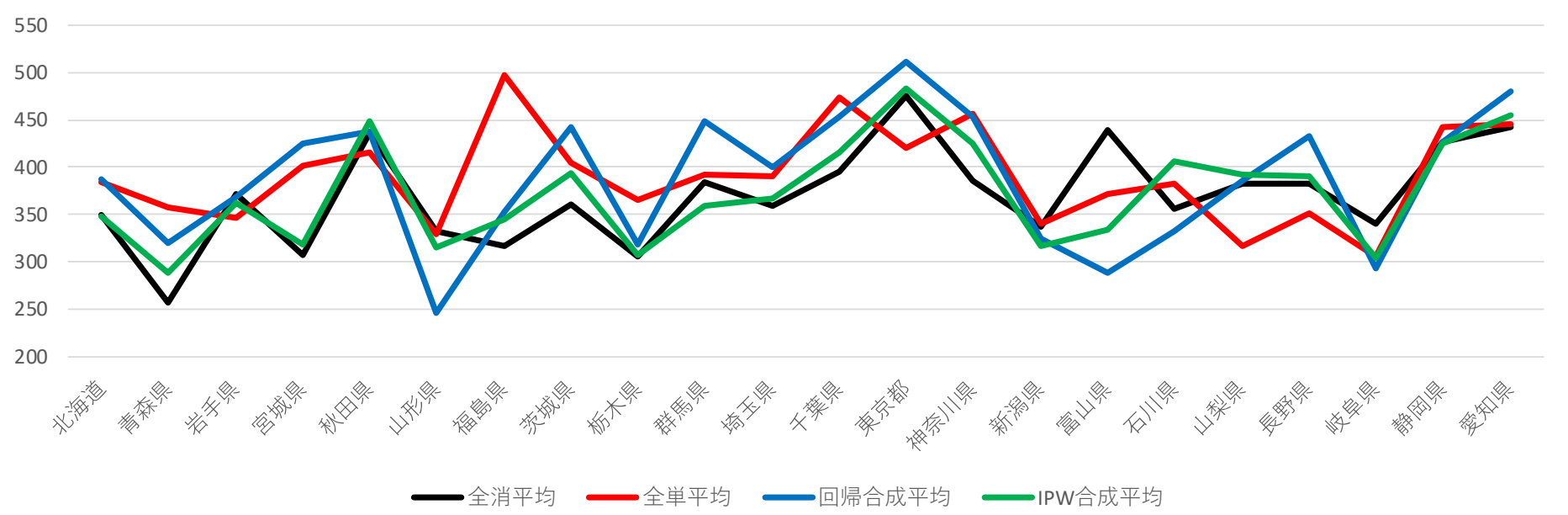
消費支出(都道府県別) 平均 回帰できたもの



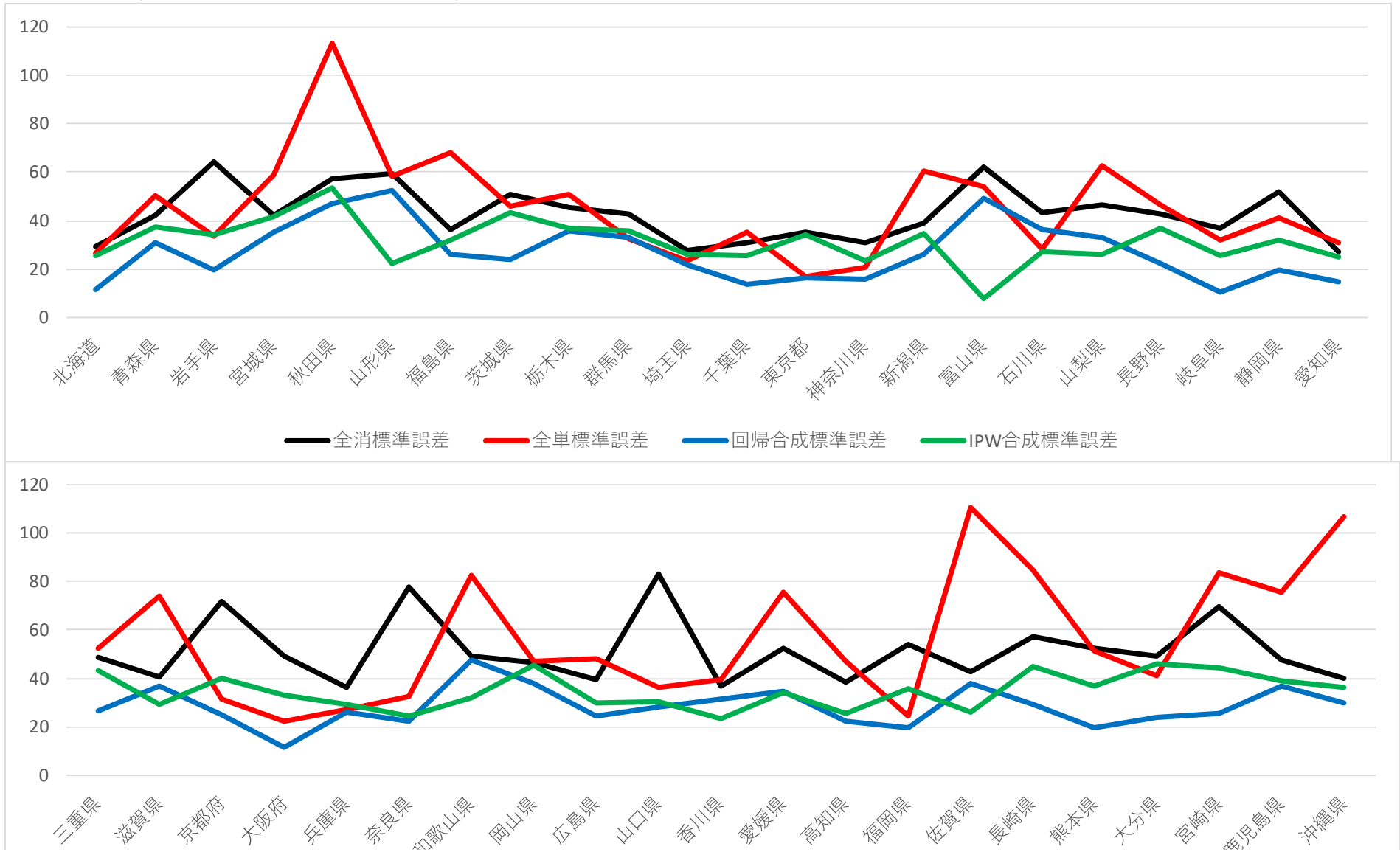
消費支出(都道府県別) 標準誤差 回帰できたもの



年間収入(都道府県別) 平均 回帰できたもの



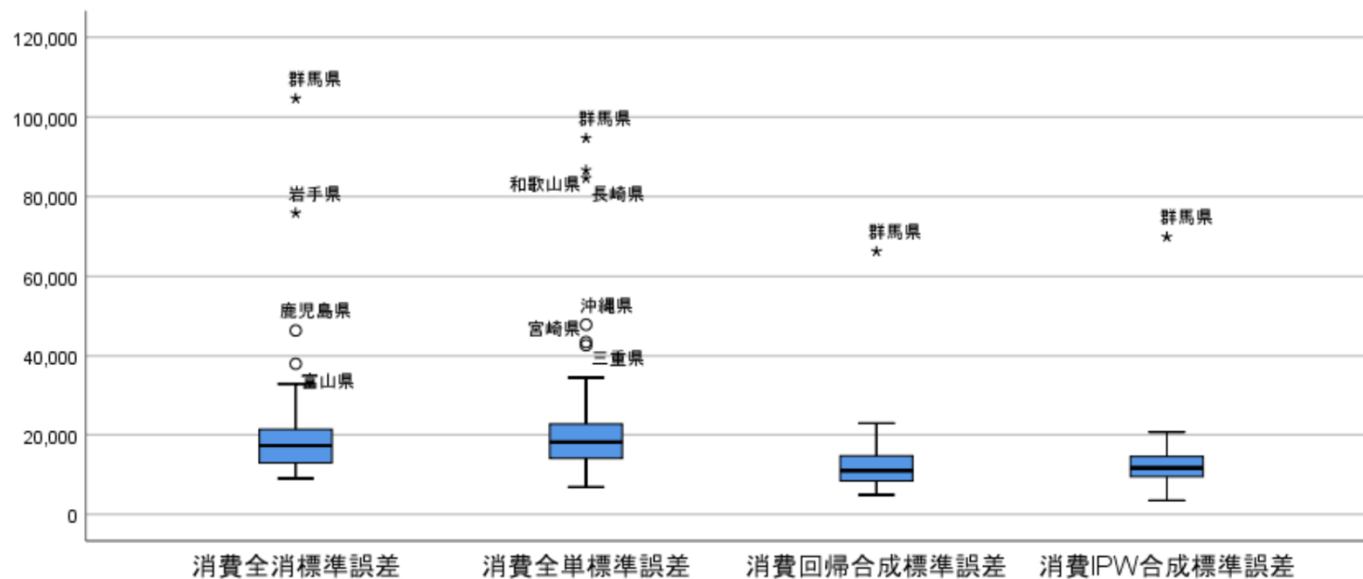
年間収入(都道府県別) 標準誤差 回帰できたもの



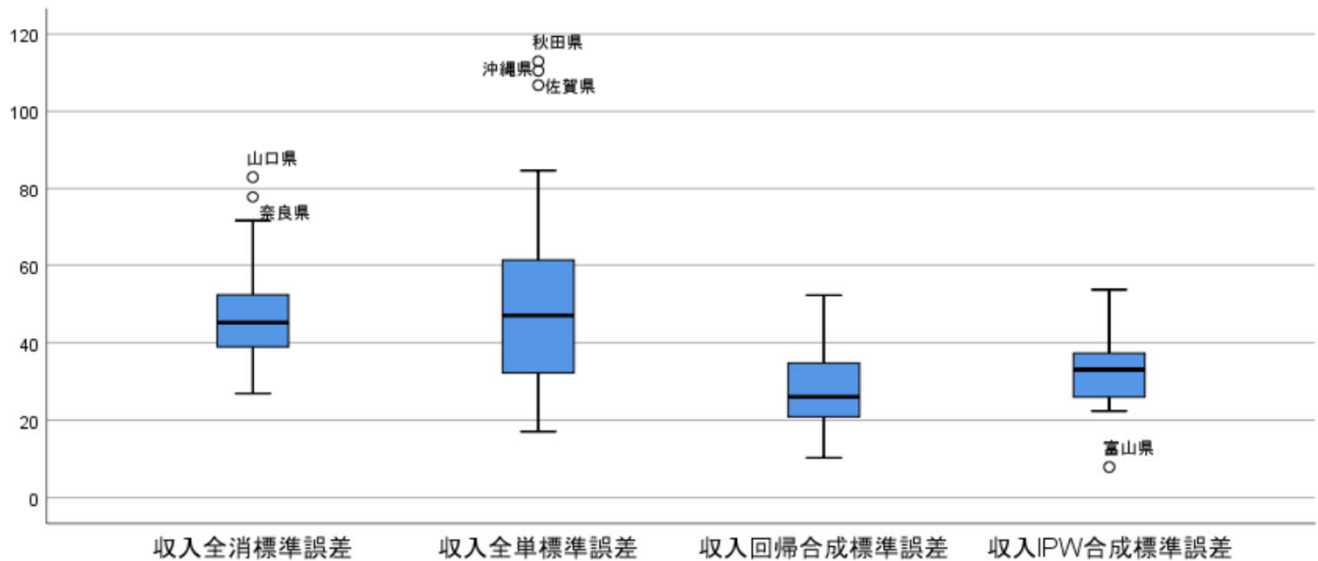
推定量の「標準誤差の分布」

回帰が計算可能なもののみ

右上：消費の平均の標準誤差の箱ひげ図



右下：収入の平均の標準誤差の箱ひげ図



消費支出(都道府県別) 回帰不能な部分

右上：平均

左上：平均の標準誤差

傾向スコアによる調整は
福井県以外では問題がない
模様

但し標準誤差については
正しく計算で来ていない
* 漸近論が成立しない
⇒ブートストラップ的な
方法で改善の可能性？



変動の原因

【平均】

消費支出

群馬・和歌山・長崎はもともと全単がかなり高い

岩手・鹿児島は全消が高い

*外れ値? 最大 (全単) 群馬217万、長崎120万 (全消) 133万 最低はそれぞれ7万

年間収入

福島・沖縄は全単が高い

富山・高知は全単が低い

【分散】 群馬の消費支出以外はほとんどサンプルサイズによる説明が可能

消費支出

岩手・群馬・和歌山・長崎⇒全単全消自体の分散が大/差異が大

秋田・沖縄⇒全単の分散が大

年間収入

秋田・滋賀・和歌山・愛媛・佐賀・宮崎・沖縄⇒全単の分散が大

まとめ

【共変量調整法の利用について】（回帰代入とIPW）

単独で利用するのではなく、全消と融合して利用

⇒推定値の分散は原則として全消より小さくなる

そのうえで安定的な推定という点ではサンプルサイズ大では回帰代入
サンプルサイズ小の場合優劣が付けづらい

【地域別の推定について】

セット2でも推定値の分散を小さくできる

【都道府県別の推定について】

推定の実行可能性という点で共変量の数に制約

そのうえで調整以前にサンプルサイズが小さい・平均の分散が大きい

⇒推定値の分散をほとんどの場合で全消より小さくできる（全単を融合する効果）

別の発想

今回の全国家計構造調査において、全国単身世帯収支実態調査は全国のデータが早めに(各県の集計表を作成する前に)整備され利用可能になるという場合を想定

(1) 都道府県ごとの家計構造と全国の全単データを融合

(2) 都道府県ごとの家計構造とその都道府県が属する地域別の全単データを融合 (例) 青森県の家計構造 + 東北の全単

これでどのような挙動を示すかを分析

* 変数セット2 どちらでも回帰ができるので全都道府県で実施可能

手法のイメージ

“調査・データ取得モードの違い”

= 取り方の違い (今回は項目の違い)

がないと仮定すると

回答集団の違い

全消調査の標本として統合 ($z = 0$)
都道府県ごと

全単調査の標本
($z = 1$) 全国/地域



全消調査の標本
($z = 0$) 都道府県ごと

全単調査の
回答結果 y_1

全単の標本での
全単調査の結果

全消の標本での全単調査
の回答結果 (欠測)

全消調査の
回答結果 y_0

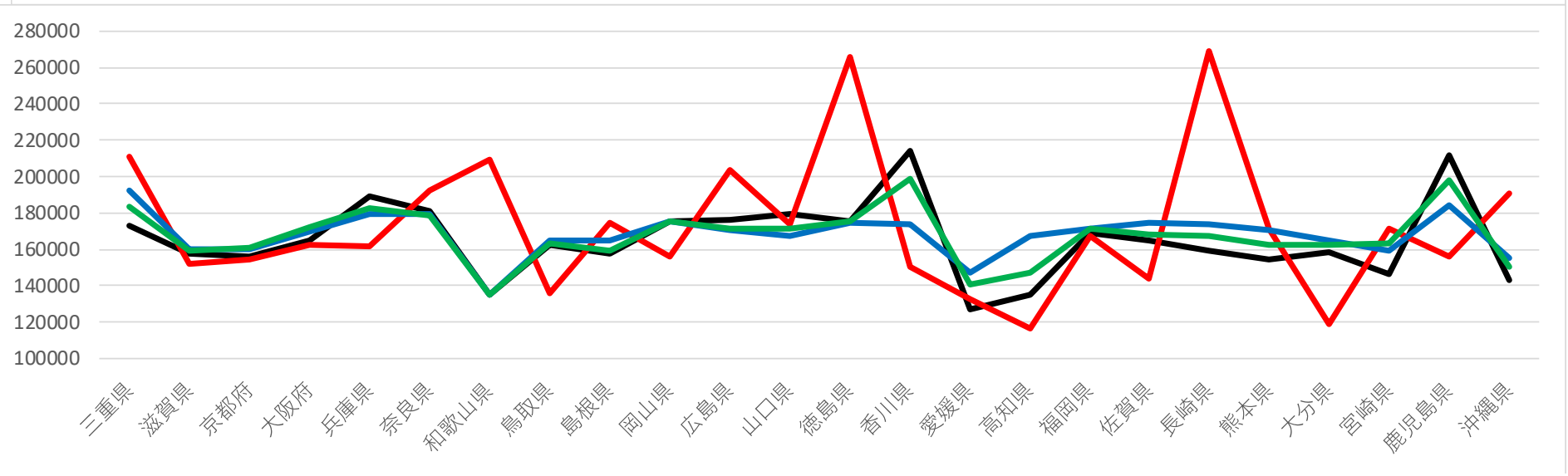
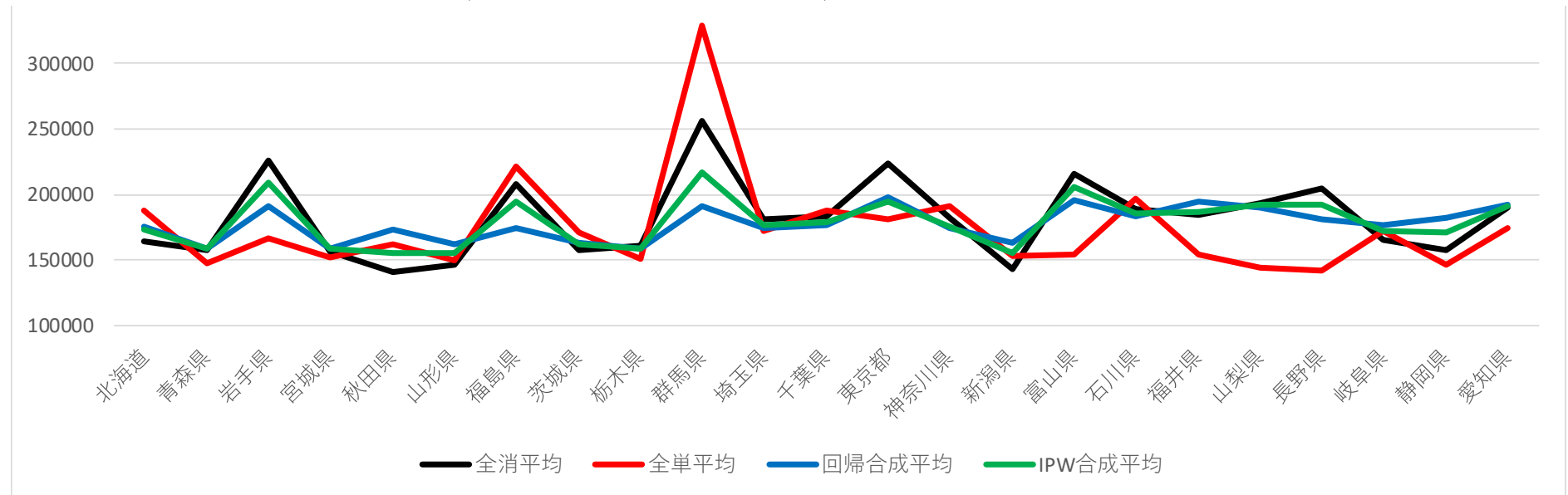
全単の標本での全消調査
の回答結果 (欠測)

全消の標本での全消調査
の結果

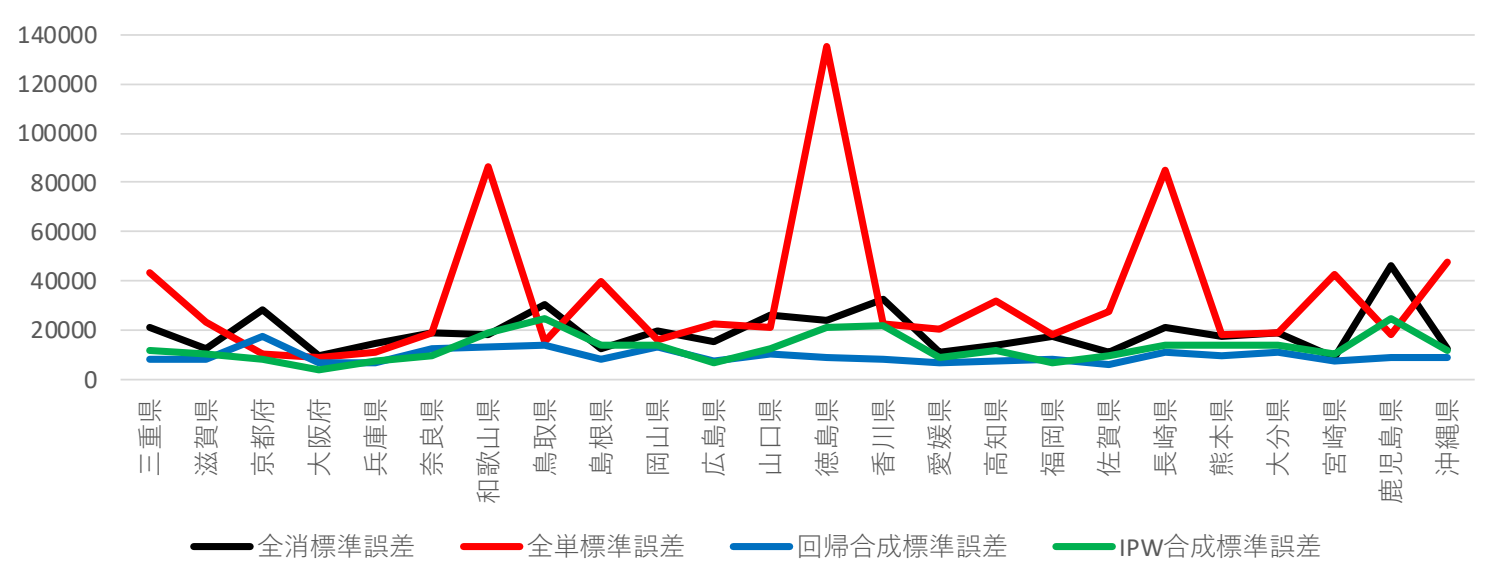
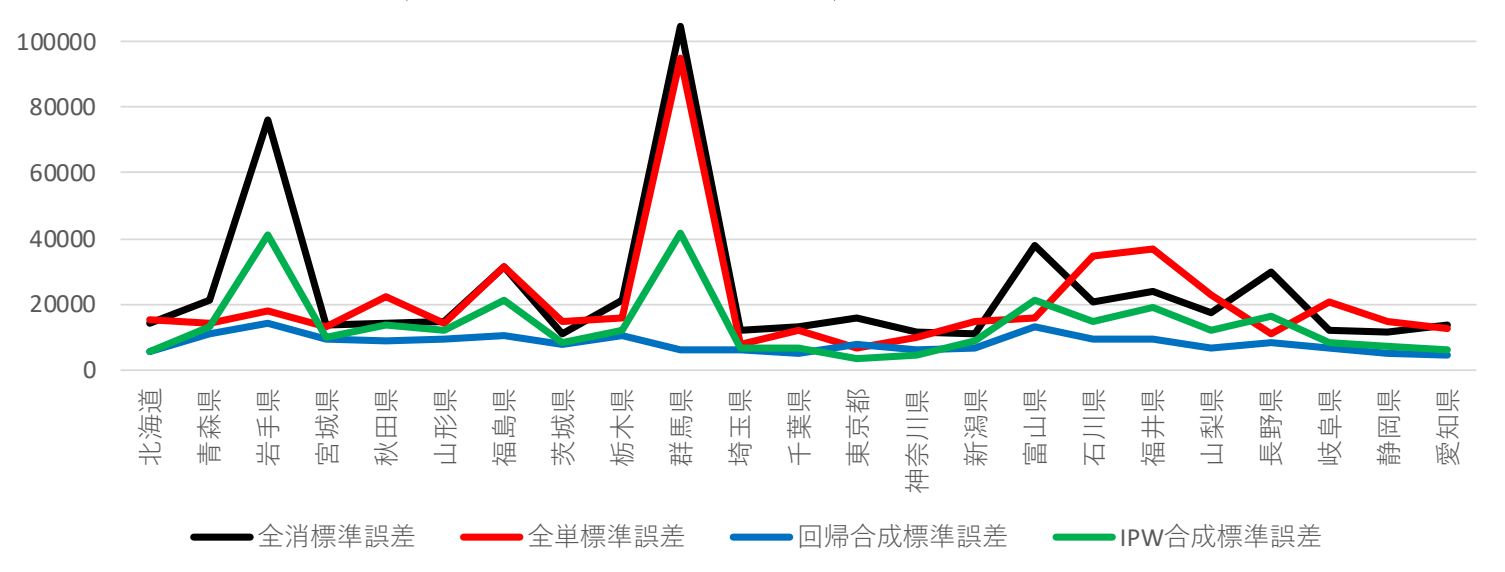
補助変数・
共変量

回答集団間の違いが生じる属性
(性年代・職種・収入等)

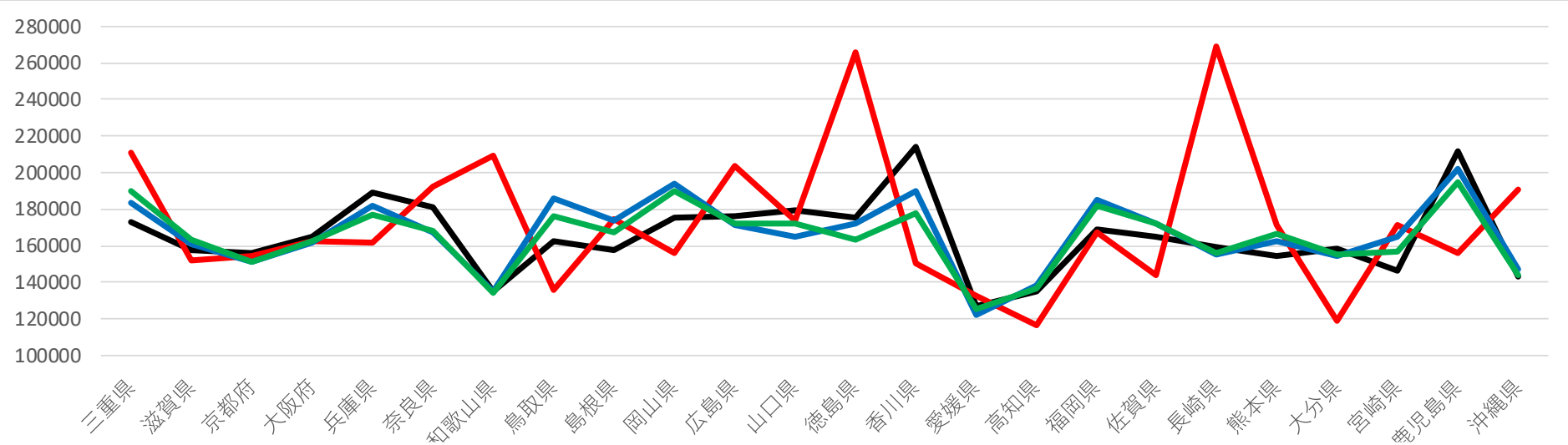
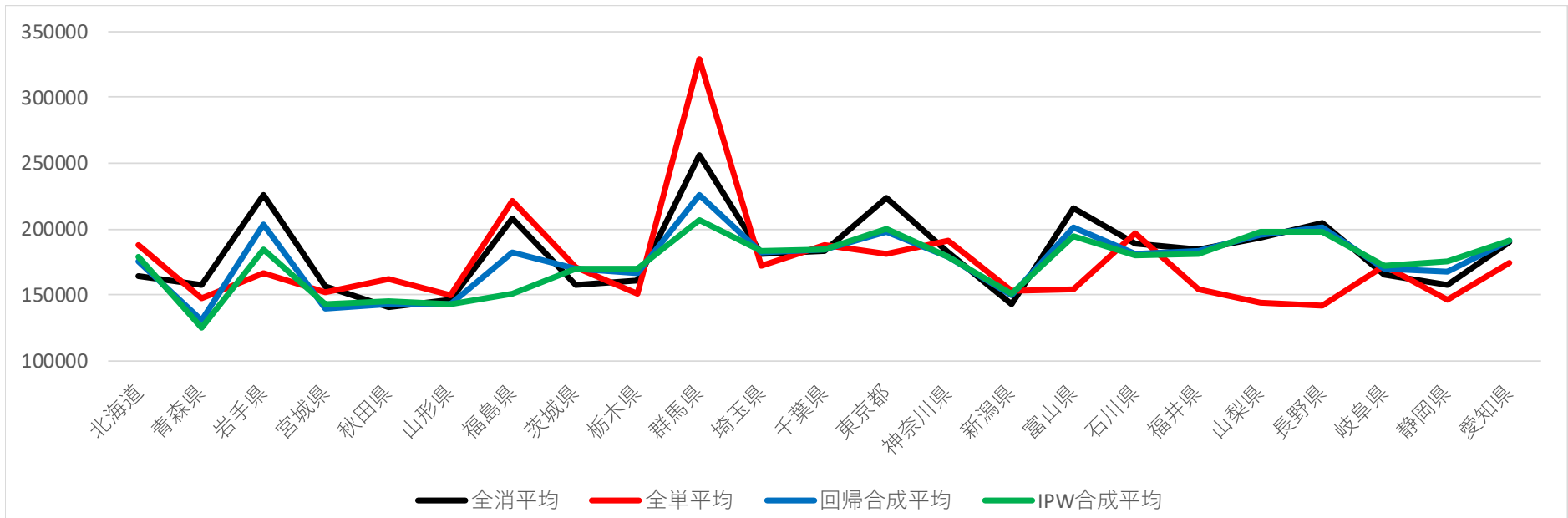
全単全国：消費支出(都道府県別) 平均



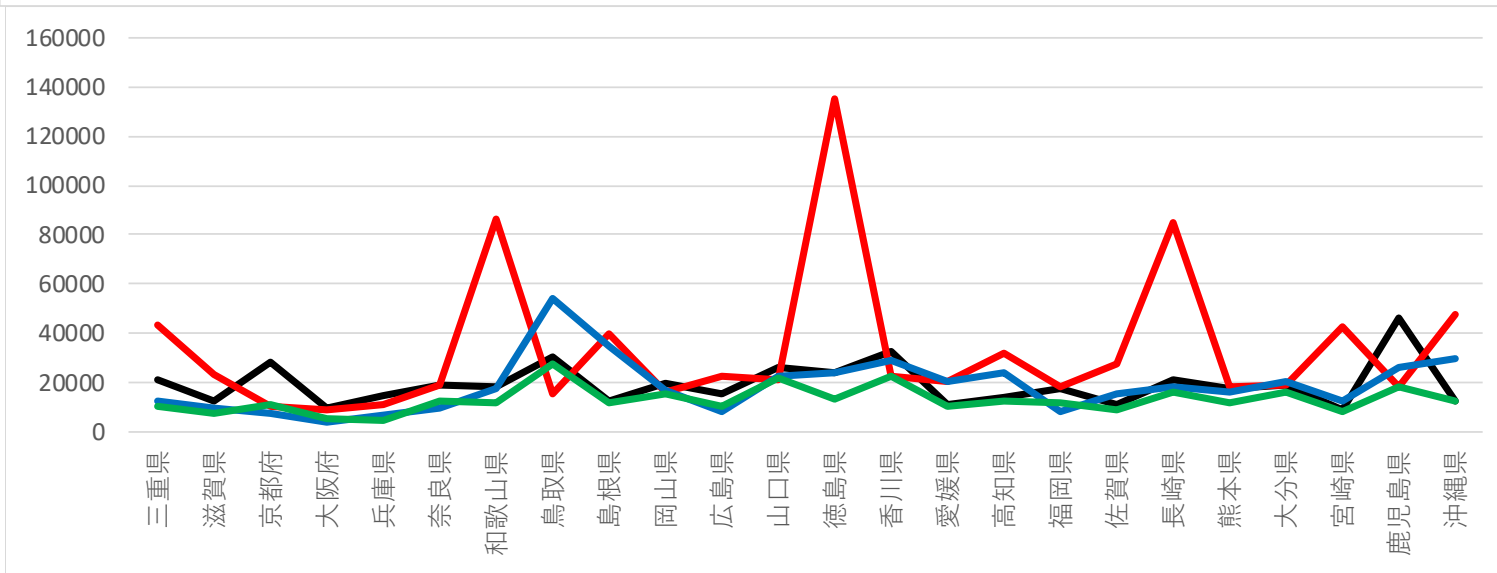
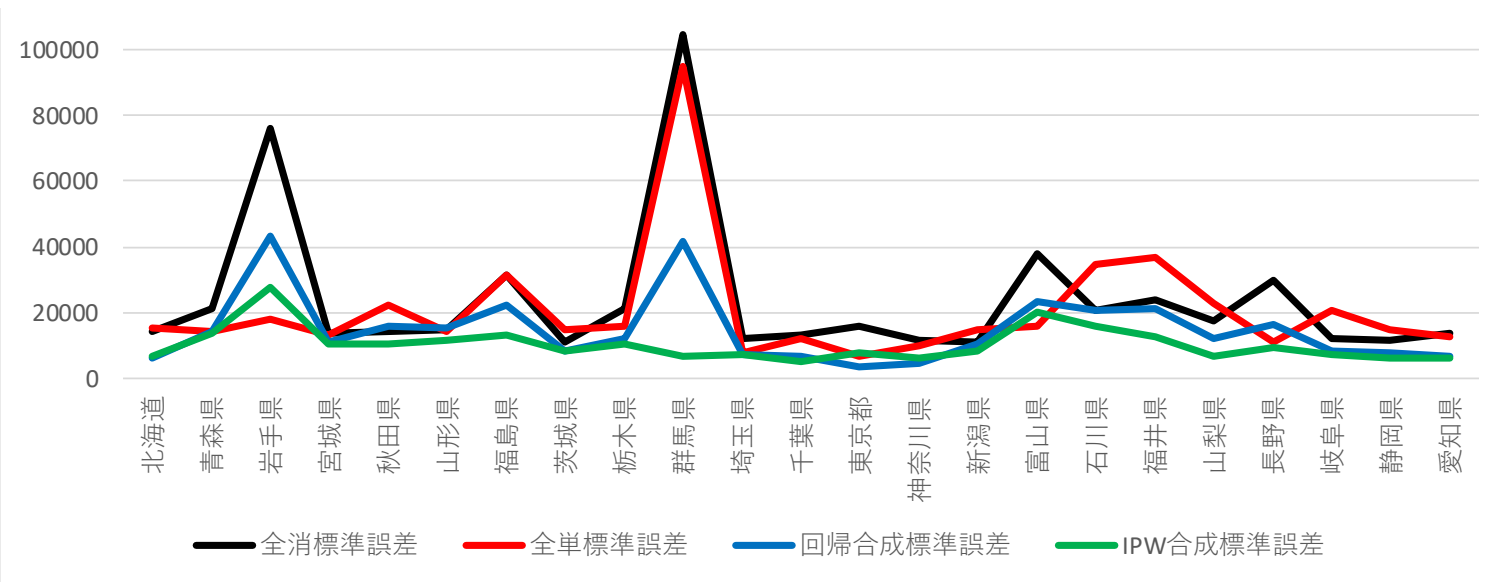
全単全国：消費支出(都道府県別) 標準誤差



全単地域：消費支出(都道府県別) 平均



全単地域：消費支出(都道府県別) 標準誤差



議論

- 8月に実施した「サンプルサイズは大きいが単身若年層の少ない全国消費実態調査（全消,今年から家計構造）と、若年層が多い全国単身世帯収支実態調査（以下,全単）のデータを融合する方法論」について、都道府県別に実施できる可能性について検討した。
- 都道府県別では全単のサンプルサイズが少なすぎるため、調整効果が大きい分散も大きくなる傾向のある傾向スコアだけでなく回帰代入法も実施し比較を行った。
- 全国／地域別では合成による精度の向上が見られた
- 都道府県別では全単のサンプルサイズが小さい県を中心に回帰分析が実行不能となるものがあった
- それ以外の都道府県については合成は一定の効果があった
⇒サンプルサイズに依存する
- 群馬県などを見ていると、特に全単について外れ値を考慮する必要性？

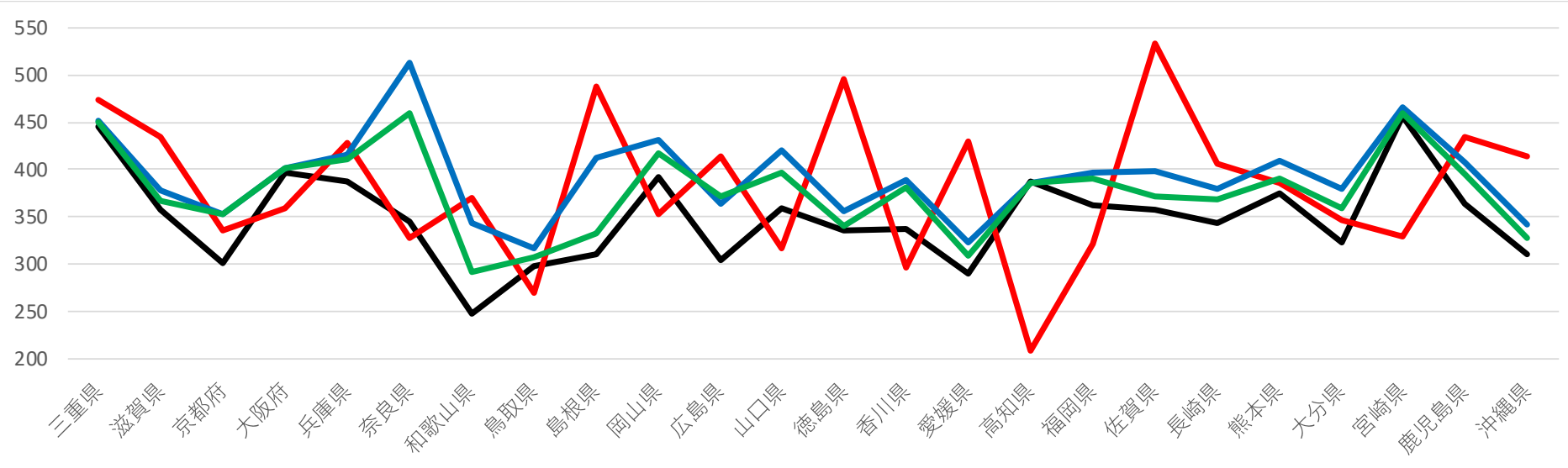
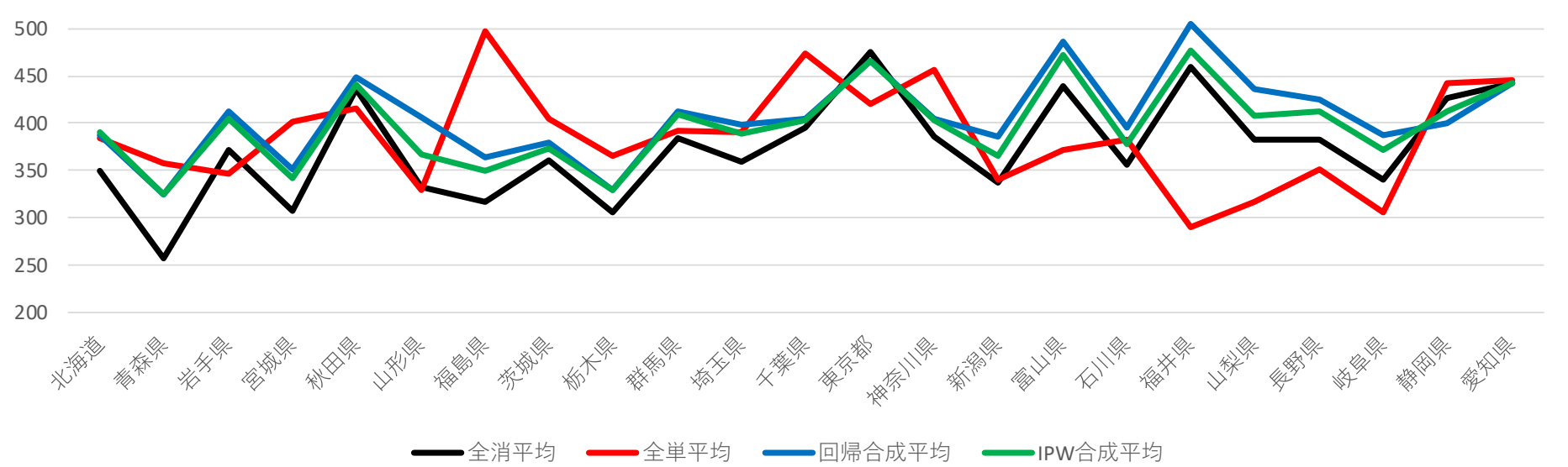
今後の方向性

- 今年の**全国家計構造**調査への全単の利用の折には全単については全国データまたは地域別データを利用することを検討することが良いかもしれない。
 - * この場合より適合度の高いモデルを利用することは可能
 - * またマッチング的な手法も利用できる(予測平均マッチングなど)
- ⇒今年度の全単のデータがいつどのようなタイミングで利用可能かに依存
都道府県別の集計表を作成する前なら有効な方法？
- 可能であれば少なくとも地域単位でデータを利用する、小地域推定の手法を利用するなど改善が行えるかもしれない。
- 精度を高めるには二重にロバストな推定法を利用(今回は実行できず)
- より詳細なレベルで許諾された研究者や事業者を利用させるなら
多重代入法の発想が必要
 - * 今回は集計表を公開するだけを検討。代入値の利用なら別。

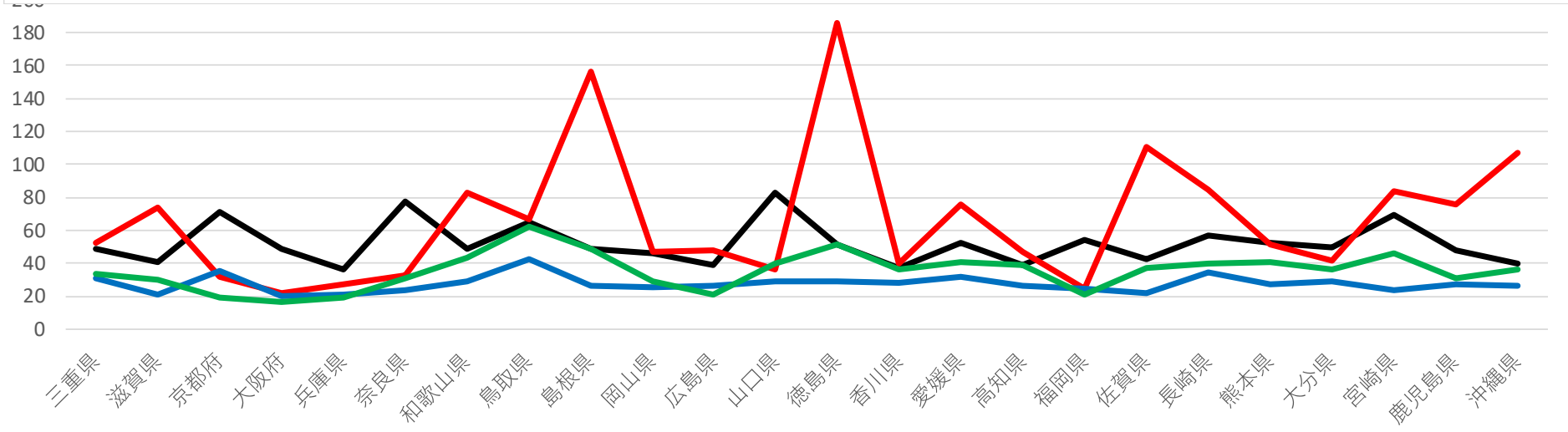
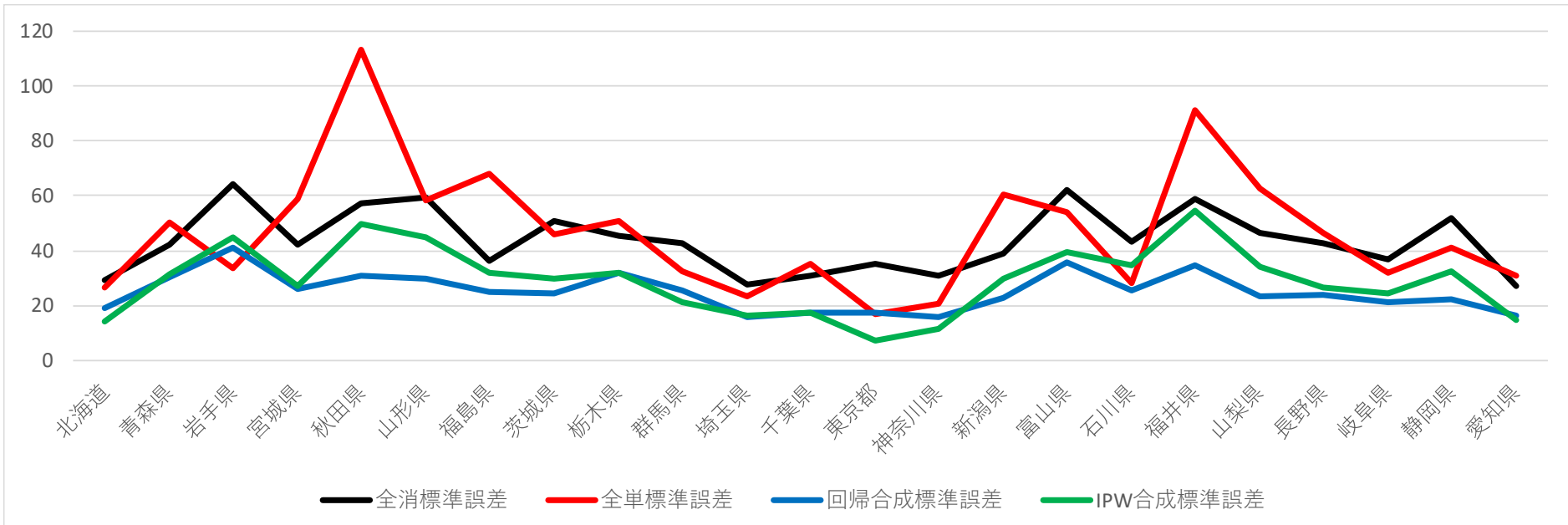
過去の両手法の利用状況

ランダムな欠測の仮定(Missing at random)

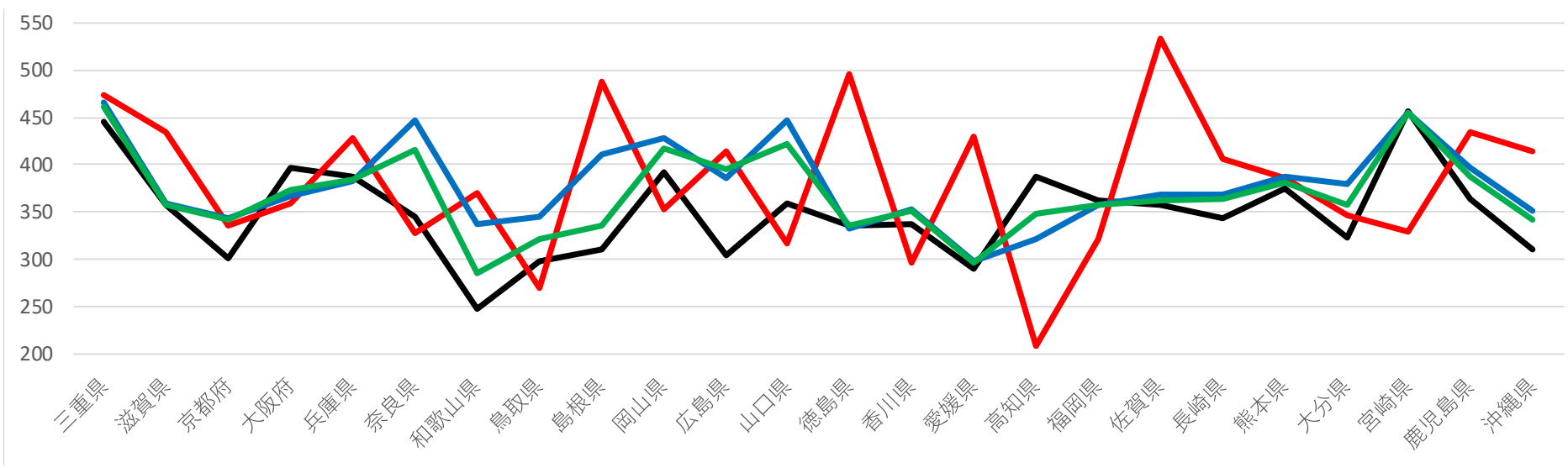
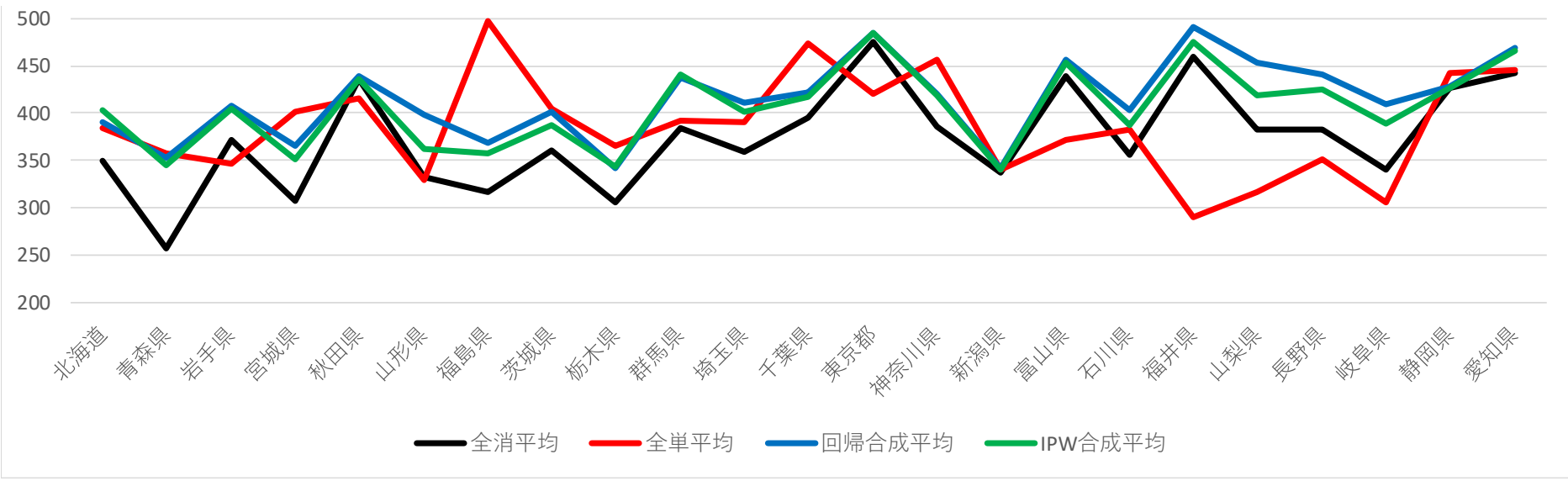
全単全国：年間収入(都道府県別) 平均



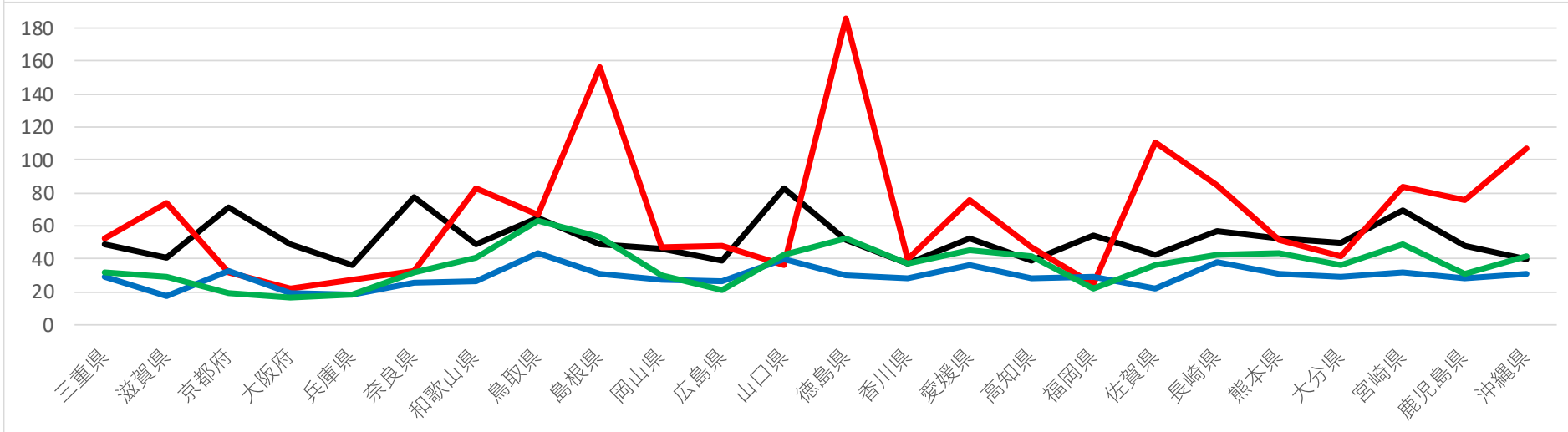
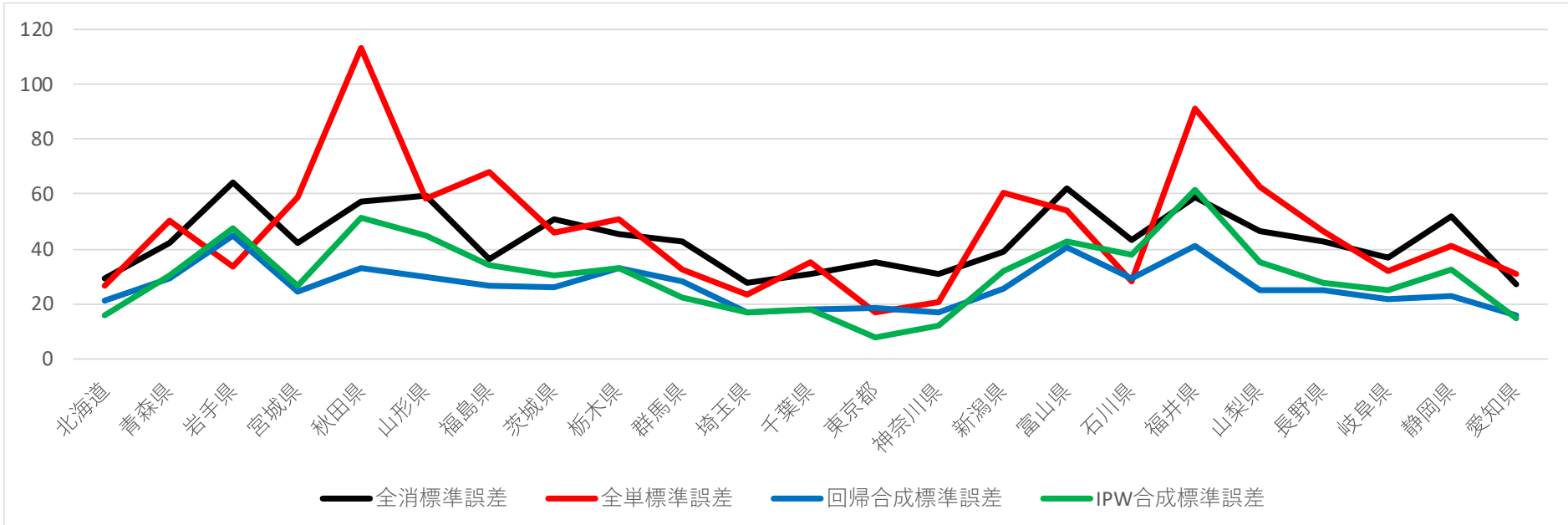
全単全国：年間収入(都道府県別) 標準誤差



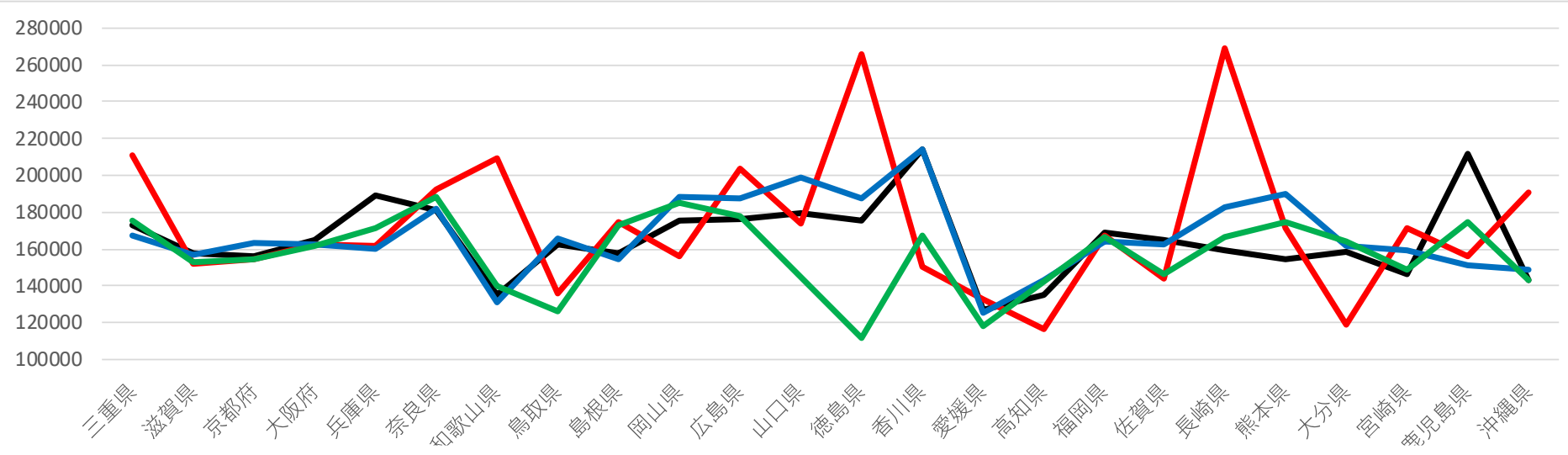
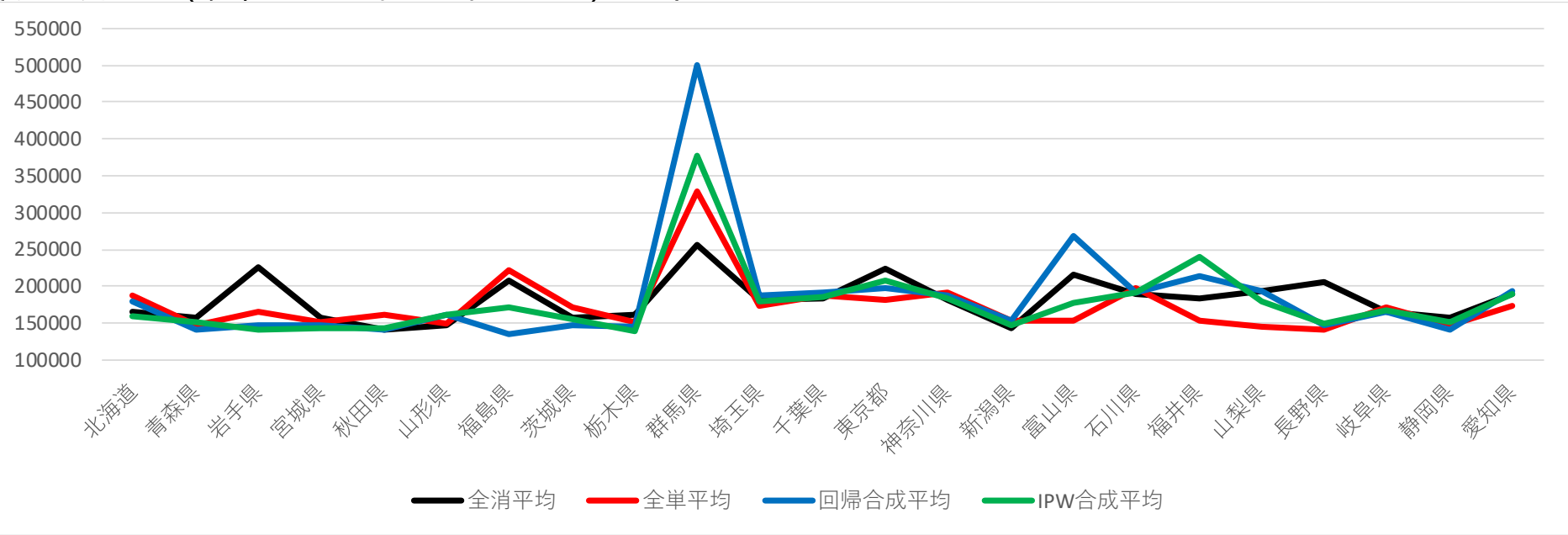
全単地域：年間収入(都道府県別) 平均



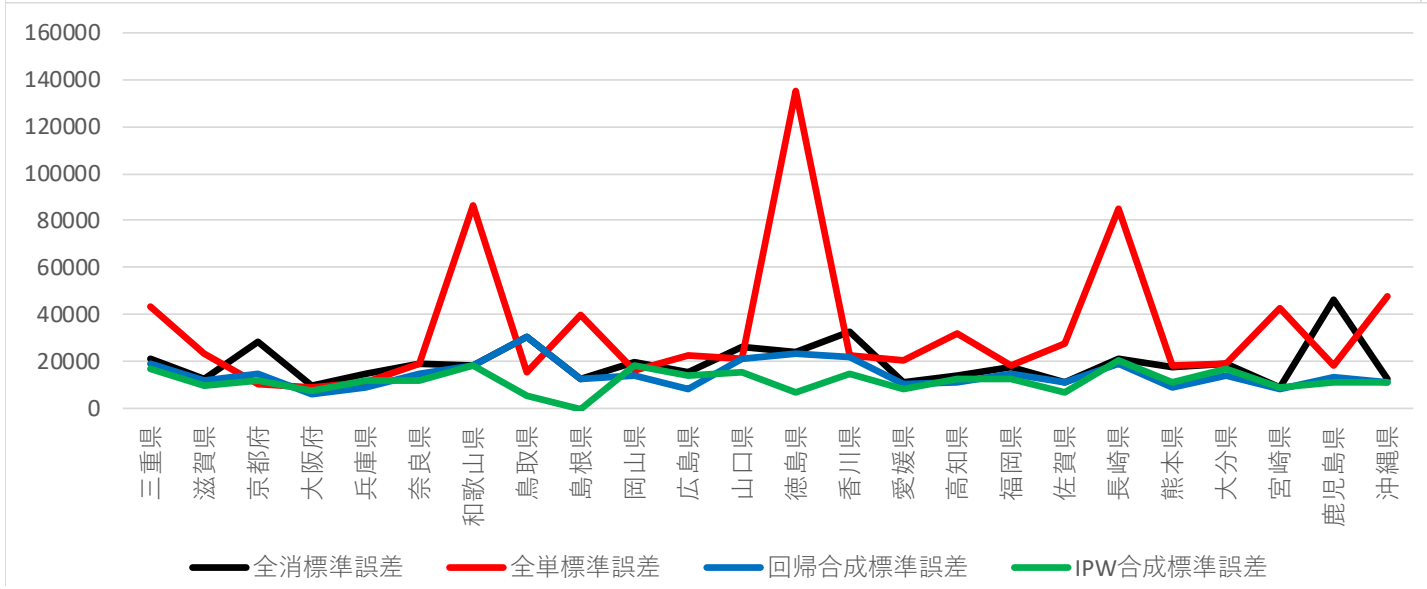
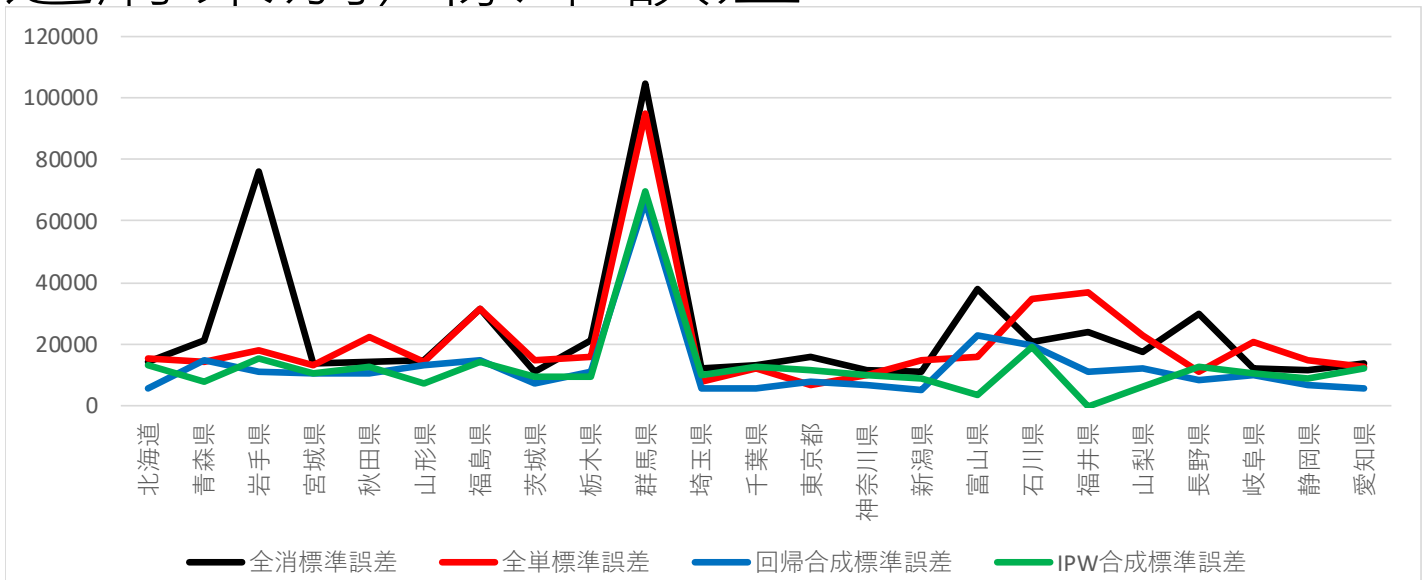
全単地域：年間収入(都道府県別) 標準誤差



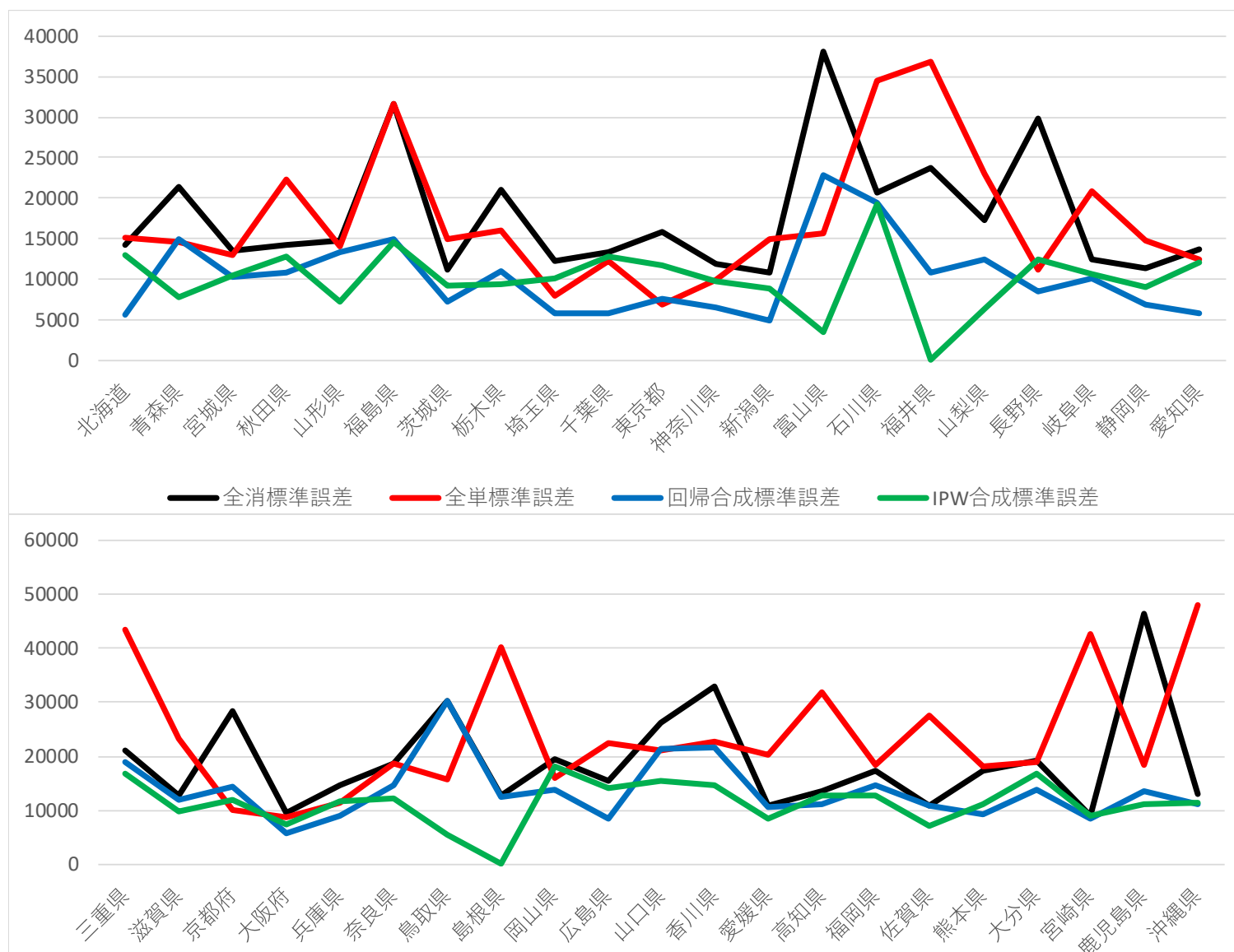
消費支出(都道府県別) 平均



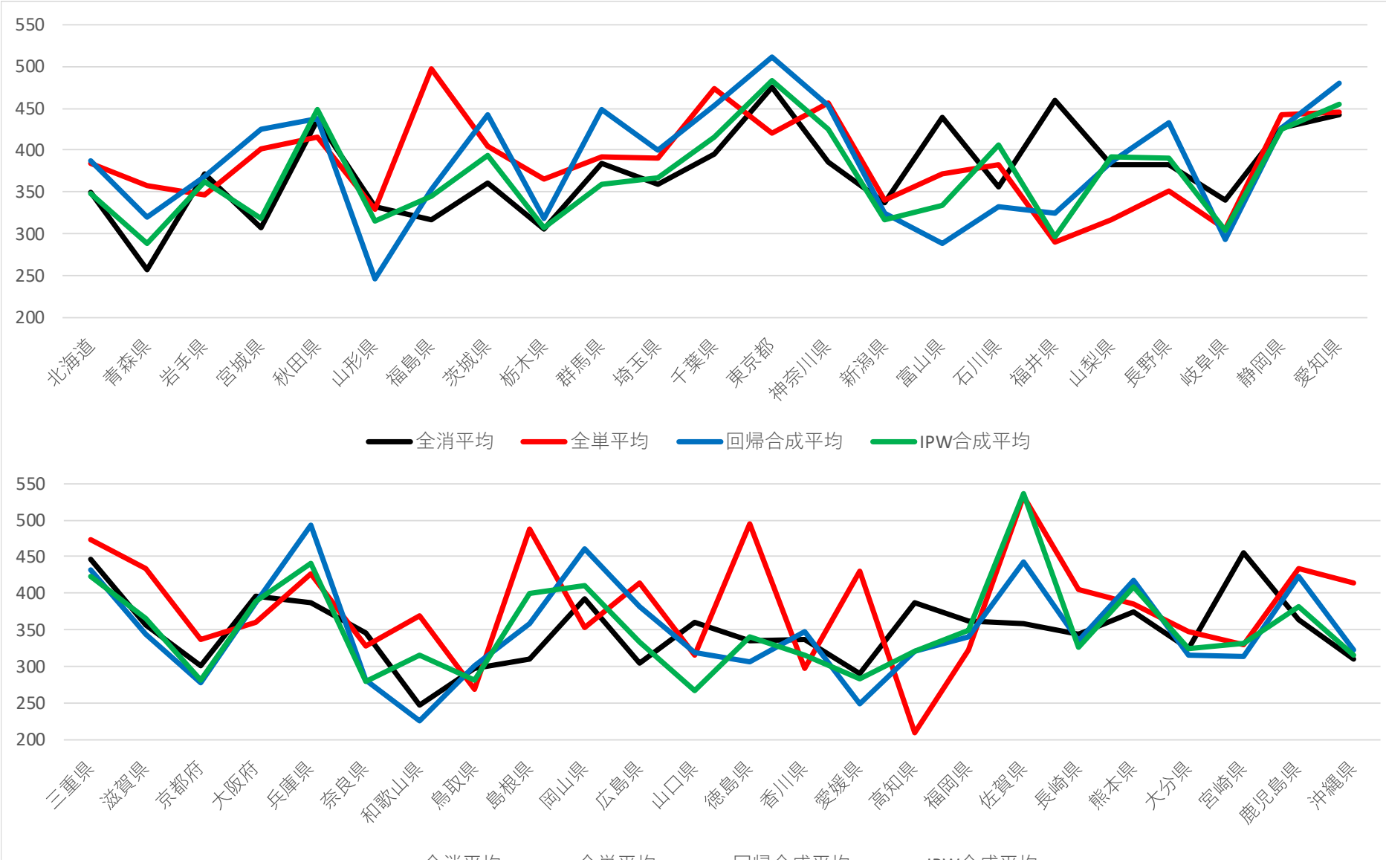
消費支出(都道府県別) 標準誤差



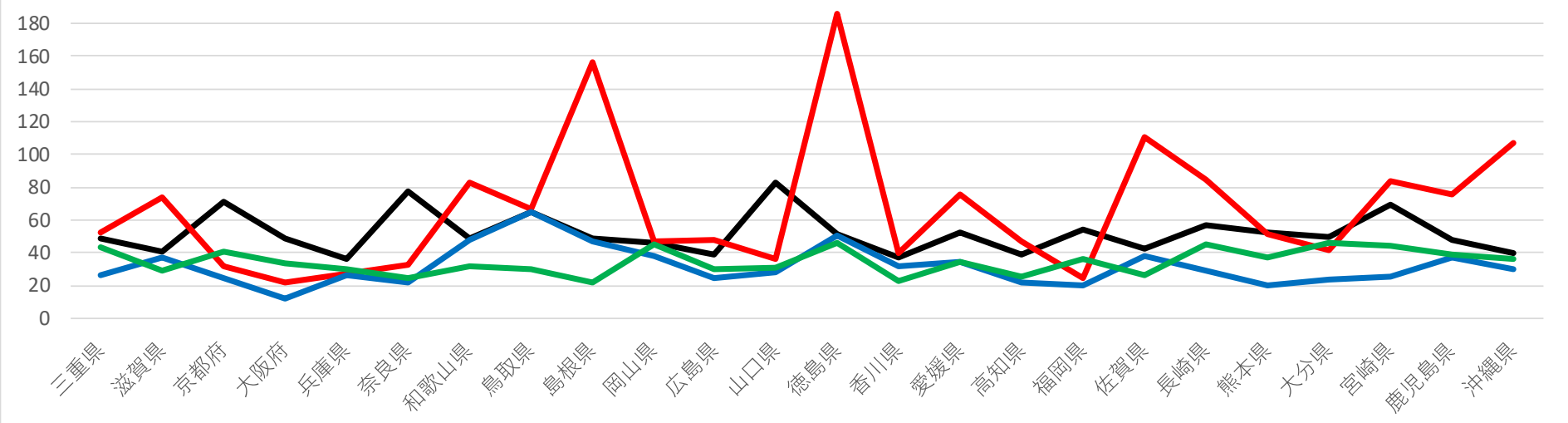
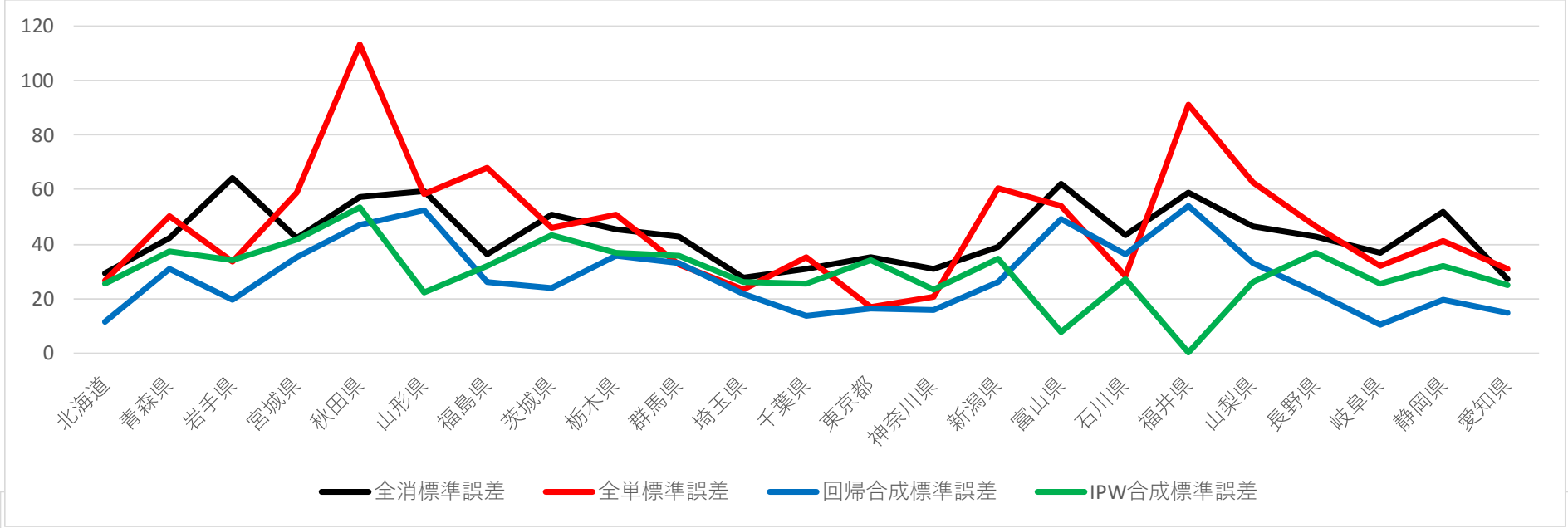
消費支出(都道府県別) 標準誤差(岩手群馬和歌山徳島長崎除外)



年間収入(都道府県別) 平均



年間収入(都道府県別) 標準誤差



年間収入(都道府県別) 標準誤差(秋田福井島根徳島佐賀削除)

