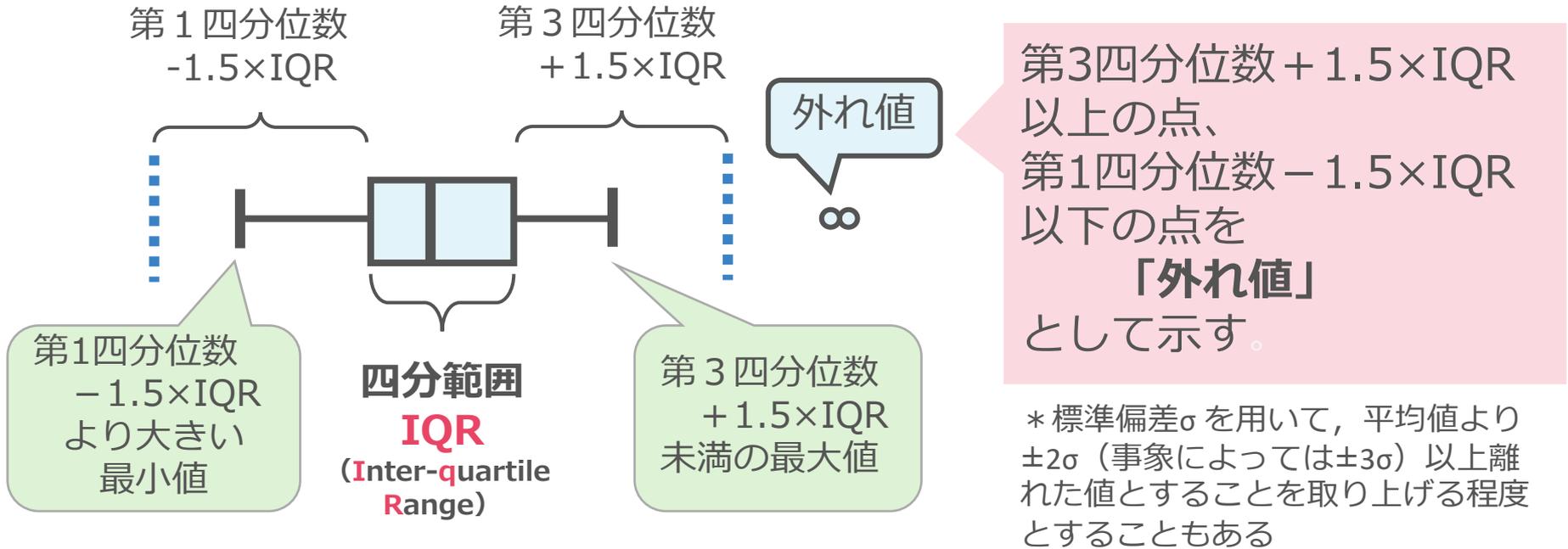


記述統計①

外れ値の意味とその活用例

独立行政法人統計センター／総務省統計研究研修所
特任講師 山下雅代

外れ値：他の値から極端にかけ離れたデータ [1]



除外した方がいい外れ値

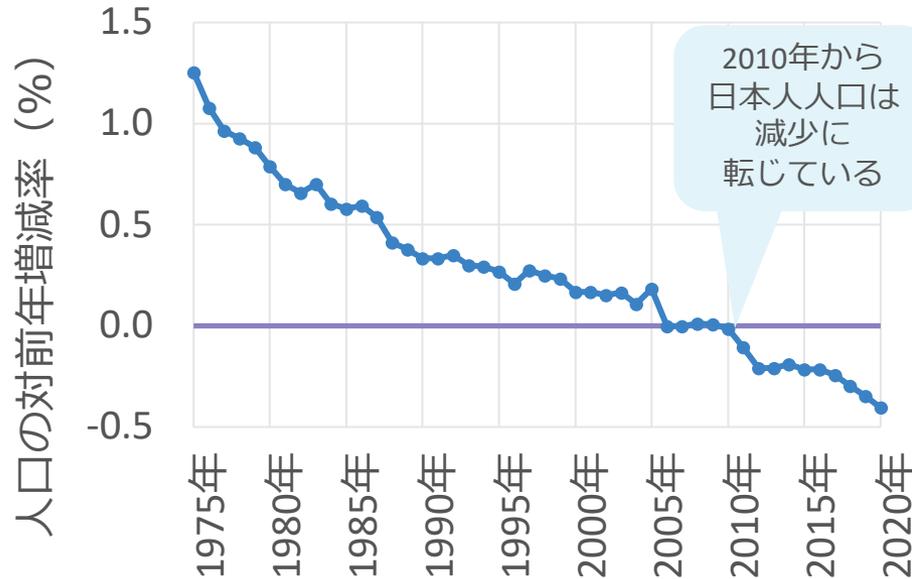
計測ミス、データの入力ミスなど → 異常値と呼ばれることも

除外しない方がいい外れ値

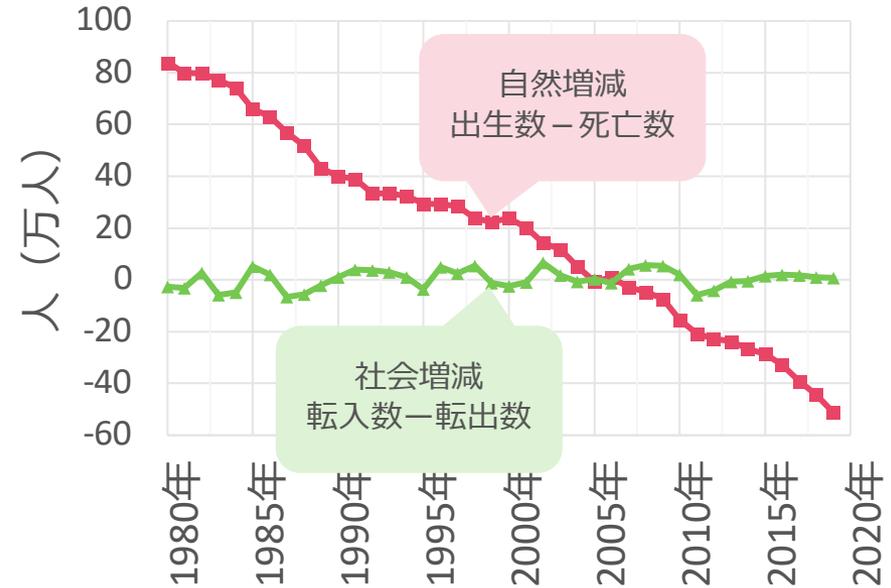
いい意味／悪い意味での外れ値 → **問題解決の着眼点**

問題場面：人口減少と少子化

日本人人口の対前年増減率



日本人の要因別の人口動態



[データの出典] 総務省 住民基本台帳に基づく人口、人口動態及び世帯数調査

出生数÷死亡数[2]を指標に用いて、

人口減少について現状把握を行い、

人口を増やすためのヒントを得ることにした

使用するデータセット

SSDSE-A-2021：市区町村別、多分野データ

SSDSE：Standardized Statistical Data Set for Education (教育用標準データセット)

日本全国1741市区町村× 125項目 の表形式

(総務省統計局「統計でみる都道府県・市区町村のすがた」から項目を抜粋、編集)

【人口・世帯】

総人口（合計、男、女）
 日本人人口（合計、男、女）
 15歳未満人口（合計、男、女）
 15～64歳人口（合計、男、女）
 65歳以上人口（合計、男、女）
 75歳以上人口（合計、男、女）
 外国人人口
 出生数、死亡数
 転入者数、転出者数
 世帯数
 一般世帯数、一般世帯人員数
 核家族世帯数
 単独世帯数
 65歳以上の世帯員のいる核家族世帯数
 高齢夫婦のみの世帯数
 高齢単身世帯数（65歳以上）
 婚姻件数、離婚件数

【自然環境】

総面積（北方地域・竹島を除く）
 可住地面積

【経済基盤】

事業所数（民営）総数
 事業所数（民営）（産業大分類別（18））
 従業者数（民営）総数
 従業者数（民営）（産業大分類別（18））
 農家数（販売農家・自給的農家）

【行政基盤】

経常収支比率（市町村財政）
 実質公債費比率（市町村財政）
 歳入決算総額（市町村財政）
 地方税（市町村財政）
 歳出決算総額（市町村財政）
 民生費（市町村財政）
 土木費（市町村財政）
 教育費（市町村財政）
 災害復旧費（市町村財政）

【教育】

幼稚園数
 幼稚園在園者数
 小学校数
 小学校教員数
 小学校児童数
 中学校数
 中学校教員数
 中学校生徒数
 義務教育学校数
 義務教育学校教員数
 義務教育学校前期課程児童数
 義務教育学校後期課程生徒数
 高等学校数
 高等学校生徒数

【労働】

就業者数（合計、男、女）
 完全失業者数（合計、男、女）
 非労働力人口（合計、男、女）
 第1次産業就業者数
 第2次産業就業者数
 第3次産業就業者数

【文化・スポーツ】

公民館数
 図書館数

【居住】

総人口（非水洗化人口+水洗化人口）
 非水洗化人口
 小売店数
 飲食店数
 大型小売店数

【健康・医療】

一般病院数
 一般診療所数
 歯科診療所数
 医師数
 歯科医師数
 薬剤師数

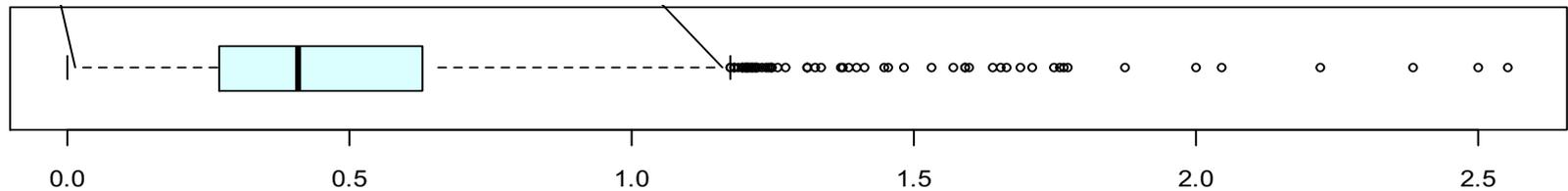
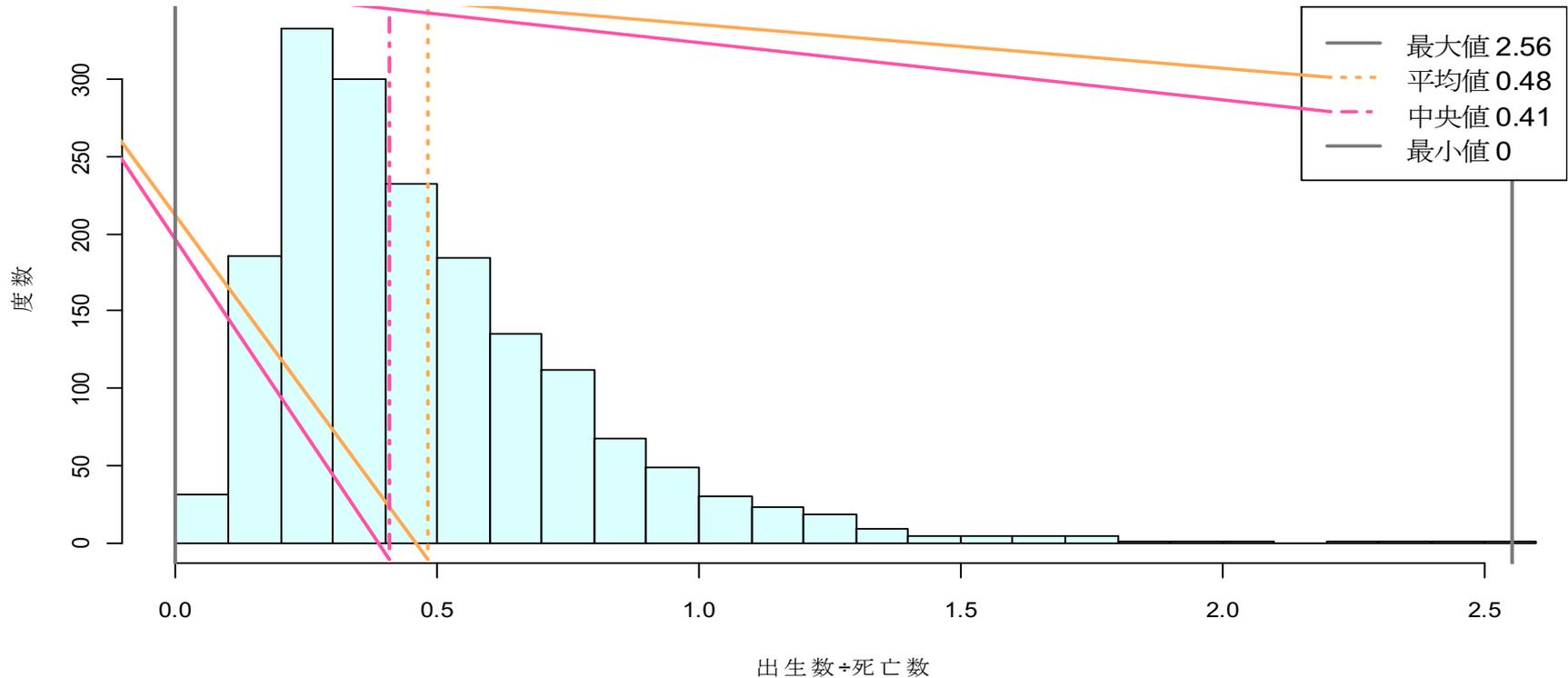
【福祉・社会保障】

保育所等数
 保育所等在所児数

日本全国1737市区町村*の「出生数÷死亡数」の分布

SSDSE-A-2021

*人口が0の4町村（福島県富岡町、大熊町、双葉町、浪江町）は除外

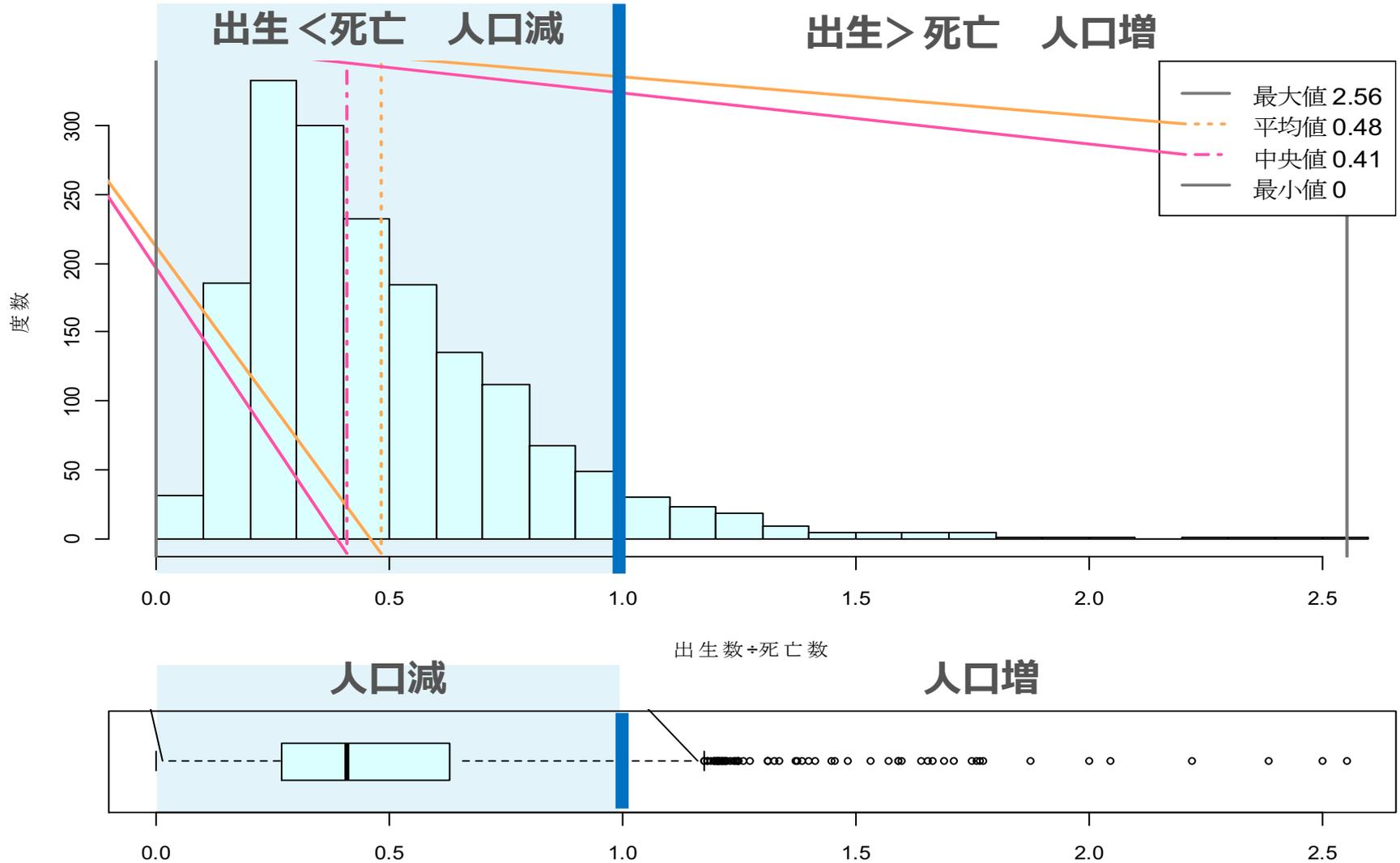


分布のどの部分が問題（人口減少）を示していますか？

日本全国1737市区町村*の「出生数÷死亡数」の分布

SSDSE-A-2021

*人口が0の4町村（福島県富岡町、大熊町、双葉町、浪江町）は除外

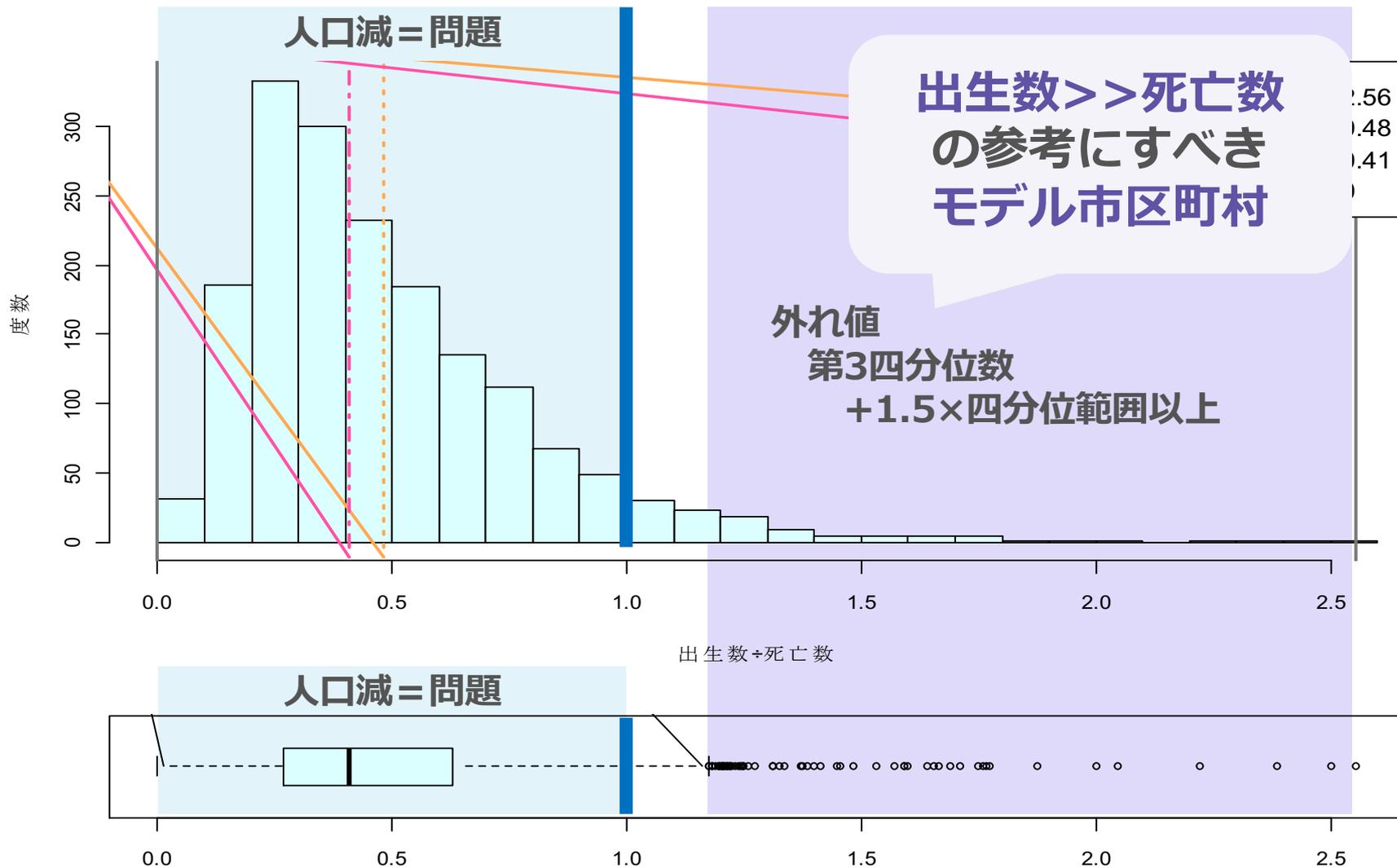


出生数÷死亡数が1未満であることが主要な問題

日本全国1737市区町村*の「出生数÷死亡数」の分布

SSDSE-A-2021

*人口が0の4市区町村（福島県富岡町、大熊町、双葉町、浪江町）は除外



外れ値はできるだけ理由を明らかにすることが大切

出生数÷死亡数の外れ値の61市区町村（SSDSE-A-2021）

都道府県	市区町村	出生÷死亡	都道府県	市区町村	出生÷死亡	都道府県	市区町村	出生÷死亡
茨城県	つくば市	1.23	愛知県	刈谷市	1.31	長崎県	佐々町	1.20
群馬県	吉岡町	1.26	愛知県	安城市	1.20	鹿児島県	三島村	1.67
埼玉県	戸田市	1.37	愛知県	東海市	1.18	沖縄県	宜野湾市	1.69
埼玉県	朝霞市	1.27	愛知県	大府市	1.40	沖縄県	石垣市	1.33
埼玉県	和光市	1.59	愛知県	知立市	1.22	沖縄県	浦添市	1.53
千葉県	流山市	1.32	愛知県	日進市	1.45	沖縄県	名護市	1.25
千葉県	浦安市	1.25	愛知県	みよし市	1.46	沖縄県	糸満市	1.38
東京都	千代田区	1.64	愛知県	長久手市	2.05	沖縄県	沖縄市	1.20
東京都	中央区	2.17	愛知県	大口町	1.24	沖縄県	国分城市	1.77
東京都	港区	1.77	愛知県	幸田町	1.62	沖縄県	うるま市	1.21
東京都	文京区	1.19	三重県	朝日町	1.21	沖縄県	北谷町	1.24
東京都	台東区	1.18	東京都	東区	1.37	沖縄県	北谷町	1.18
東京都	台東区	1.18	東京都	台東区	1.59	沖縄県	北谷町	1.59
東京都	利島村	1.20	広島県	海田町	1.31	沖縄県	西原町	1.38
東京都	青ヶ島村	2.00	福岡県	大野城市	1.18	沖縄県	与那原町	1.76
東京都	小笠原村	2.56	福岡県	宮町	1.65	沖縄県	南風原町	2.38
石川県	野々市市	1.37	福岡県	粕屋町	1.71	沖縄県	座間味村	1.22
山梨県	昭和町	1.41	福岡県	上野市	1.20	沖縄県	北大東村	2.50
岐阜県	瑞穂市	1.24	熊本県	合志市	1.18	沖縄県	八重瀬町	1.31
岐阜県	幸野町	1.43	熊本県	大津町	1.19	沖縄県	竹富町	1.60

これら一つ一つについて
その特徴を調べるのは大変...

データを要約する

代表値による比較を考える

沖縄県、首都圏、愛知県が多い

相関分析：結果 y に対する原因 x の探究

結果 y = 出生数、
原因 x の候補に、出生数を除く、SSDSE-Aの124項目を使用して
相関分析を行う

【人口・世帯】

総人口（合計、男、女）
日本人人口（合計、男、女）
15歳未満人口（合計、男、女）
15～64歳人口（合計、男、女）
65歳以上人口（合計、男、女）
75歳以上人口（合計、男、女）
外国人人口
死亡数
転入者数、転出者数
世帯数
一般世帯数、一般世帯人員数
核家族世帯数
単独世帯数
65歳以上の世帯員のいる核家族世帯数
高齢夫婦のみの世帯数
高齢単身世帯数（65歳以上）
婚姻件数、離婚件数

【自然環境】

総面積（北方地域・竹島を除く）
可住地面積

【経済基盤】

事業所数（民営）総数
事業所数（民営）（産業大分類別（18））
従業者数（民営）総数
従業者数（民営）（産業大分類別（18））
農家数（販売農家、自給的農家）

【行政基盤】

経常収支比率（市町村財政）
実質公債費比率（市町村財政）
歳入決算総額（市町村財政）
地方税（市町村財政）
歳出決算総額（市町村財政）
民生費（市町村財政）
土木費（市町村財政）
教育費（市町村財政）
災害復旧費（市町村財政）

【教育】

幼稚園数
幼稚園在園者数
小学校数
小学校教員数
小学校児童数
中学校数
中学校教員数
中学校生徒数
義務教育学校数
義務教育学校教員数
義務教育学校前期課程児童数
義務教育学校後期課程生徒数
高等学校数
高等学校生徒数

【労働】

就業者数（合計、男、女）
完全失業者数（合計、男、女）
非労働力人口（合計、男、女）
第1次産業就業者数
第2次産業就業者数
第3次産業就業者数

【文化・スポーツ】

公民館数
図書館数

【居住】

総人口（非水洗化人口+水洗化人口）
非水洗化人口
小売店数
飲食店数
大型小売店数

【健康・医療】

一般病院数
一般診療所数
歯科診療所数
医師数
歯科医師数
薬剤師数

【福祉・社会保障】

保育所等数
保育所等在所児数

出生数との相関係数

出生数との相関係数（一部）

項目	相関係数
総人口	0.995
総人口（男）	0.994
総人口（女）	0.994
15歳未満人口	0.994
15歳未満人口（男）	0.994
15歳未満人口（女）	0.994
15～64歳人口	0.995
15～64歳人口（男）	0.995
15～64歳人口（女）	0.995
65歳以上人口	0.982
65歳以上人口（男）	0.982
65歳以上人口（女）	0.980
75歳以上人口	0.978
75歳以上人口（男）	0.979
75歳以上人口（女）	0.976
外国人人口	0.882

- 出生数と他の項目との相関係数を算出すると左の表のような結果が出る

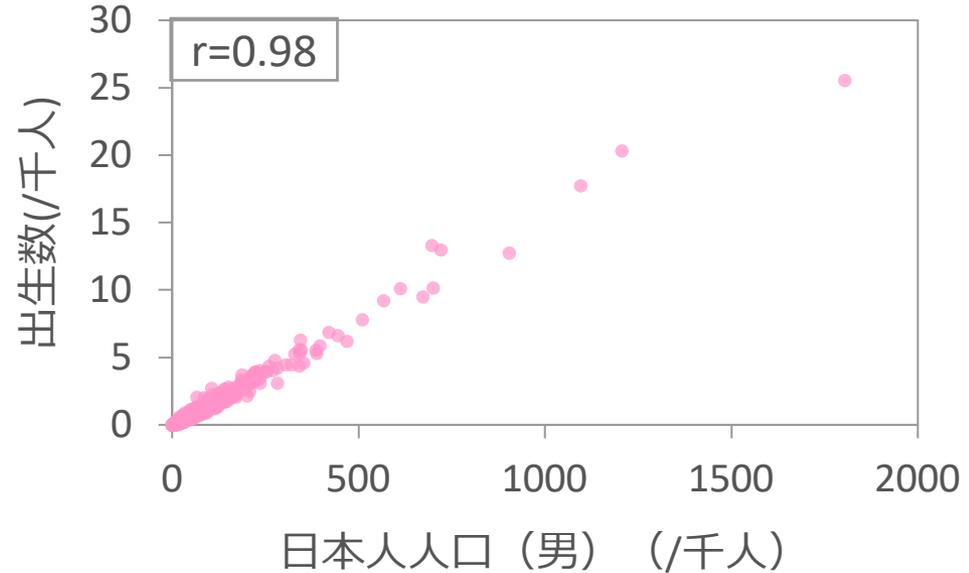
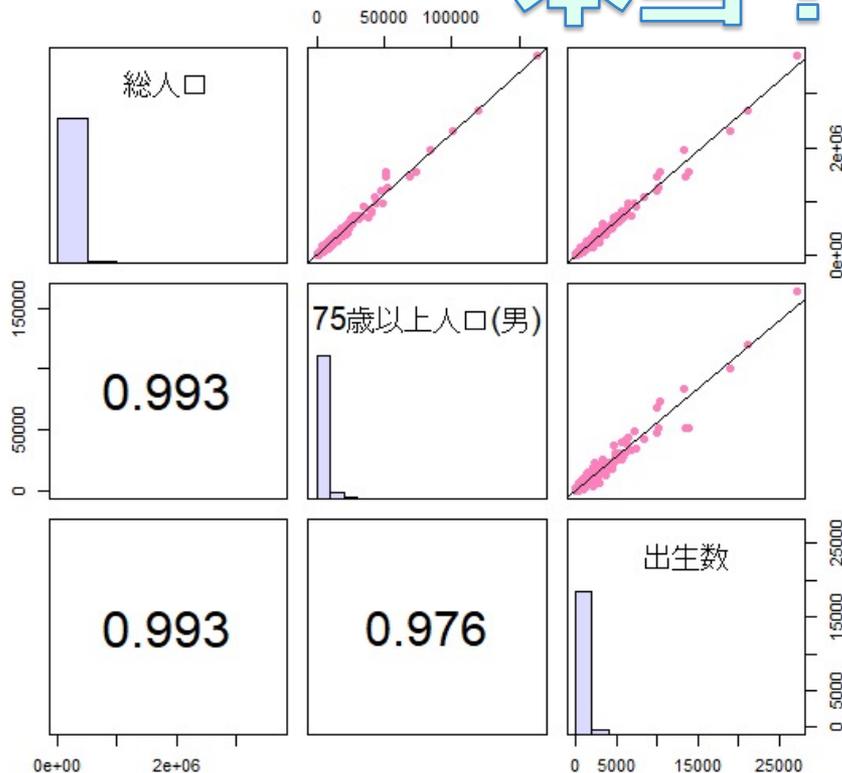
なぜ多くの項目で非常に高い相関係数が観察されるのか？

出生数と75歳以上男性数の関係

出生数と75歳以上男性数の
相関係数は0.98である

75歳以上の男性数が多いと、
子供がたくさん生まれる

本当？



総人口と75歳以上男性数 : $r = 0.993$

総人口と出生数 : $r = 0.993$

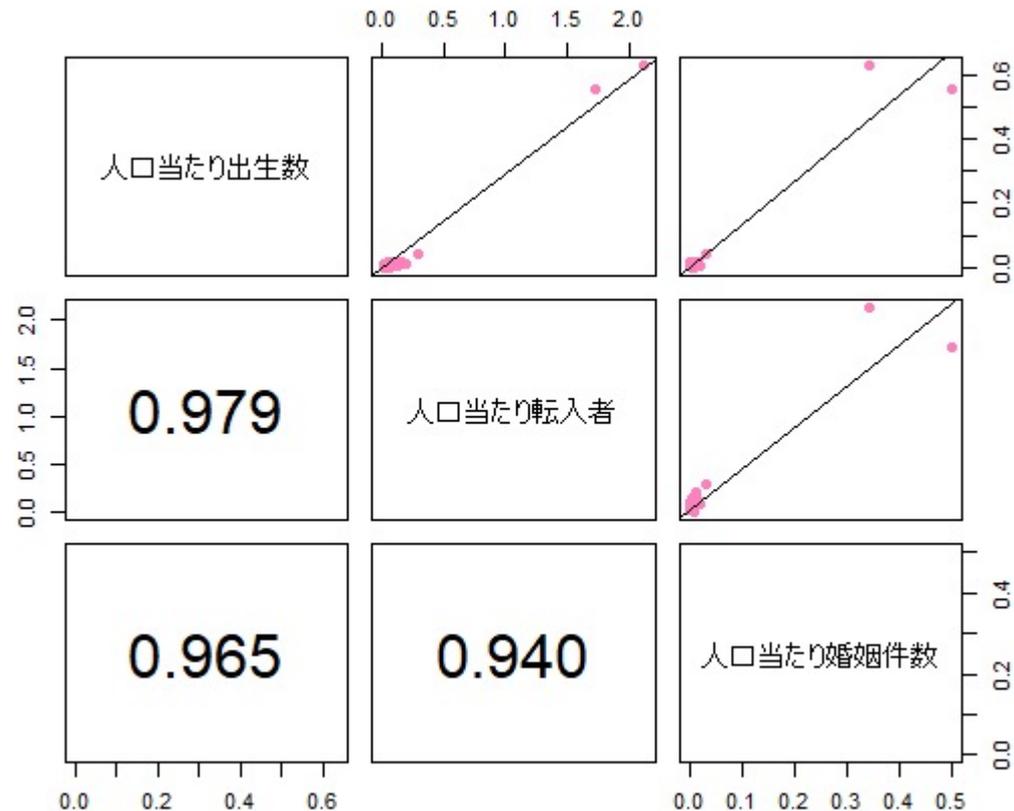
⇒ 総人口を介して75歳以上男性数と
出生数が関係があるかのように
みえてしまう

→ 疑似相関

総人口の影響を除くために
各変数を総人口で割る必要がある

人口当たり出生数と各変数の相関係数

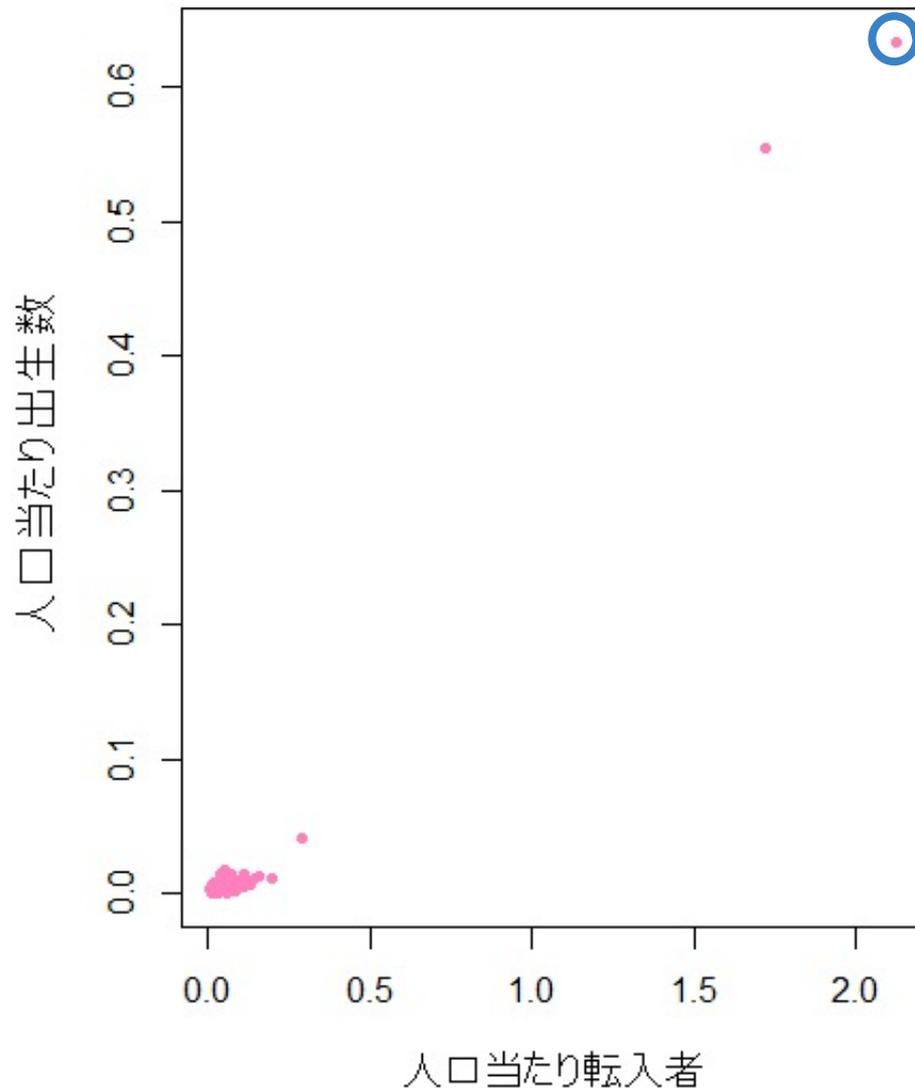
人口当たり出生数と 相関の高い項目	相関係数
歳入決算総額.市町村財政.	0.993
転入者数.日本人移動者.	0.979
小学校数	0.969
婚姻件数	0.965
可住地面積	0.958
地方税.市町村財政.	0.944
死亡数	0.927
実質公債費比率.市町村財政.	0.921
離婚件数	0.910
転出者数.日本人移動者.	0.904
民生費.市町村財政.	0.893



質問

これらの散布図の問題は？

外れ値のある散布図



福島県飯館村

- 人口 : 41人
- 出生数 : 26人
- 転入者数 : 87人
- 人口は国勢調査2015年データ
- 2017年9月5日一部を除いて避難指示解除

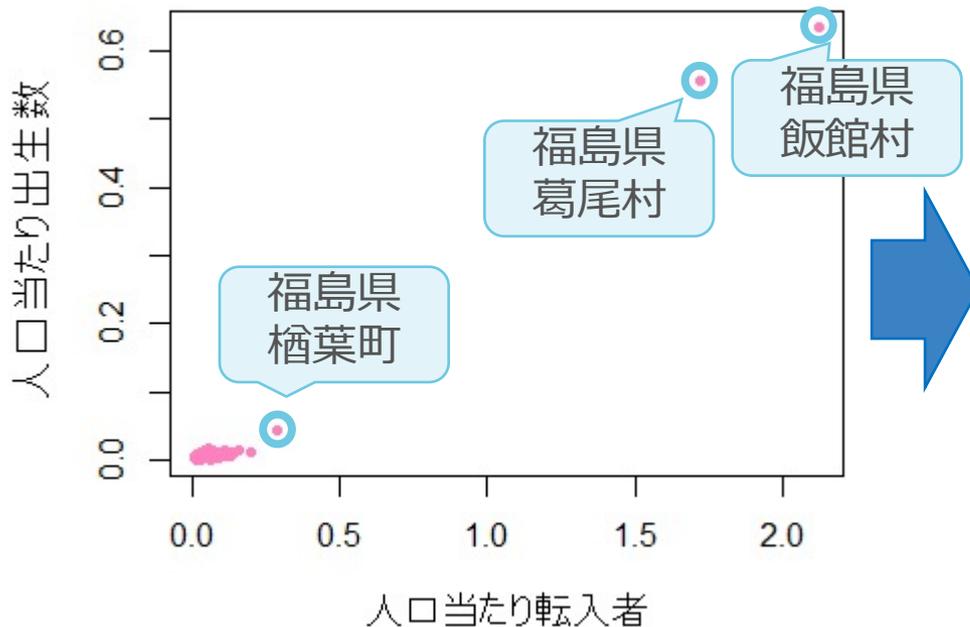


分析の趣旨に適さない

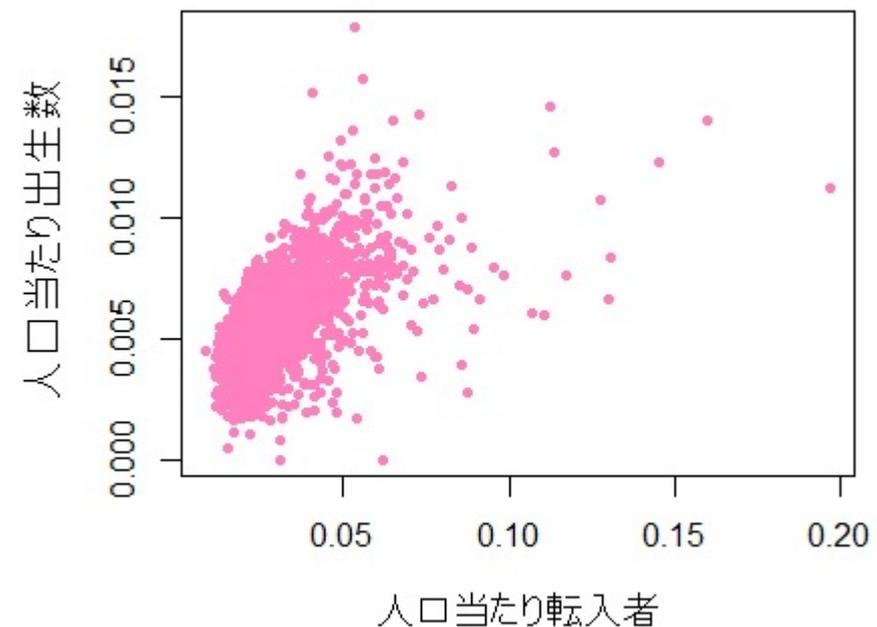
除去すべき外れ値

転入件数と出生数の散布図（一人当たり）

相関係数 $r = 0.98$



相関係数 $r = 0.58$



外れ値

除外した方がいい場合・除外しない方がいい場合

外れている理由を調べることが大切

SSDSE統計教材 日本統計協会の月刊誌『統計』 [3]

「授業に使えるSSDSEの統計教材（中学・高校編）」が掲載
2019年4月号から2020年6月号まで全12回

月刊誌 『統計』	授業に使えるSSDSEの統計教材（中学・高校編）	学習内容
2019年 4月号	第1回 新学習指導要領が目指す統計教育とその教材	
5月号 6月号	第2～3回 主体的・対話的な学びを実現する箱ひげ図の教材	箱ひげ図 外れ値
5月号 6月号	第4～5回 外れ値に着目しよう –分布の可視化と問題発見–	外れ値 ヒストグラム
7月号 8月号	第6～7回 2変数間の関係性を観察しよう –相関分析の二つの落とし穴–	散布図、外れ値 疑似相関
9月号 10月号	第8～9回 必要な汗はかこう –インターネットを用いた情報収集と地域特性の分類–	周辺情報の収集 と分析
2020年 2月号 4月号	第10～11回 2019年度版SSDSEの公開と統計地図作成	GIS、散らばり 方と分類
6月号	第12回 教材の振り返りと実社会の問題解決の方法	問題解決サイク ル

SSDSE（教育用標準データセット）

- 統計センターが作成・公表、**完全フリー**
- 欠測値がなく、すぐに分析に使えます
- 外れ値の理由を調べることができます

DLは
こちらから



SSDSEの公開状況（2021年8月現在）

名称	SSDSE-A-2021	SSDSE-B-2021	SSDSE-C-2021	SSDSE-D
内容	市区町村別、 多分野データ	都道府県別、時系列、 多分野データ	都道府県庁所在市別、 家計消費データ	都道府県別、自由時間 活動・生活時間データ
表のサイズ	縦 1744 × 横 128	縦 566 × 横 108	縦 50 × 横 229	縦 146 × 横 122
地域数	1741市区町村	47都道府県	全国 + 47県庁所在市	全国 + 47都道府県
元になる 統計	総務省「統計でみる都道府県・市区町村の すがた（社会・人口統計体系）」		総務省「家計調査」	総務省 「社会生活基本調査」
項目数	125	105	226	119
年次	各項目、 最新の1年次	2007～2018年 （12年次）	2018年～2020年の 平均値	2016年
提供開始年	2018年	2019年	2020年	2021年

※ 毎年、最新データに更新、今後も内容の充実を図る（新たな SSDSE の開発・提供など）

参考文献

- [1] 文部科学省(2018)：『高等学校学習指導要領（平成30年告示）解説 数学編
理数編』， https://www.mext.go.jp/content/1407073_05_1_2.pdf
- [2] 日本品質管理学会監修(2017)：「データに基づく問題解決」
https://www.soumu.go.jp/main_content/000544539.pdf

<使用したデータ>

- 独立行政法人統計センター **SSDSE-A-2021**（教育用標準データセット・市区
町村別多分野データ） <https://www.nstac.go.jp/SSDSE/>
- 総務省 **住民基本台帳に基づく人口、人口動態及び世帯数調査**
https://www.soumu.go.jp/main_sosiki/jichi_gyousei/daityo/jinkou_jinkou_doutai-setaisuu.html