

## 【統計ミニ解説 1】年齢調整と標準化について、その初歩から紹介します

### 年齢調整の意義とは？

市(区)町村のデータを見比べる時にしばしば起こる話の流れが、そのままの(=粗の)数字をとりあげて『〇〇は、高齢者層が多いから…』『◇◇は、若年者層が多いから…』で終わる結論です。「その地域でその数値になるのは、住民の年齢構成から考えればそんなもの」ということだと思います。健康関連のデータは年齢が大きく影響しやすいものですから、そのとおりなのでしょう。

しかし、ある地域において対策をとるべき健康課題を明らかにする時に、年齢を調整しないデータだけを見比べて『年齢構成が違うから…』で済ませることはもったいないと言えます。たしかに、問題の実際のインパクトの大きさ(コストなどが良い例です)は粗の数値の大小としてあらわれることが多いです。あるいは、普段の業務において数字の種類が増えることをわずらわしく感じるかもしれません。しかし、年齢を調整した数値を比べることで、住民の年齢構成の違いにかかわらず存在する特徴、さらには課題に気づきやすくなります。

同時に、データを住民(対象者)全体に占める割合や、ひとりあたりの数値として表現することで、人口規模の違いにかかわらず数値を比べることができるでしょう。健康課題においてスケールメリット・人口密度の違いが影響することは少なくないですが、ひとりあたりの表現はそういった影響も見やすくしてくれます。

以上の理由からこのサイトでも、住民の年齢構成の違いと人口規模の違いに関わらず存在する特徴を見えやすくするために、「年齢調整された」「住民(対象者)全体に占める割合や、ひとりあたりの数値」のデータが多く集積・配布されています。ここでは、年齢調整の代表的な方法である標準化の意味するところについて、手順を追いながら少していねいに考えていきましょう。

### 標準化(年齢調整の間接法)について

あるデータについて年齢構成の違いを踏まえて比較するための最もシンプルな方法は、集計を年齢階級別に分けて、それぞれの年齢階級同士で数値を比較することでしょう。たとえば「40歳～44歳では…」「45歳～49歳では…」と順番に論じていきます。

ただ、この方法では市区町村の姿をひと目でとらえることは難しいでしょう。私たちにとって便利なのは、年齢構成の違いを踏まえつつ、まずは全体として比較するための数値です。そして、その数値の算出方法の1つが標準化(年齢調整の間接法)です。

標準化で求められる数値は、ある集団(ここでは、ある市(区)町村とします)と基準集団(県全体とします)の比、すなわち“標準化XX比”として表現されます(XXは着目しているデータの事象)。このサイトでは多くの場合、県全体=100と設定した時の市(区)町村の値です。

### 標準化の手順を追ってみる ——国保医療費(被保険者ひとりあたり総点数)を例に

例として、平成28年度のKDBデータより医療費に着目した“標準化医療費比”をとりあげます(注:医療費の高低を、そのまま良い・悪いに結び付けていないことにご注意ください)。元データは、KDB帳票No.8(S21\_009)『医療費の状況』から抽出した年齢別「レセプト総点数(医科)」 「被保険者数」と、これらより計算された「被保険者1人あたり総点数」です。なお、計算の詳細は文末の《補足》に整理しました。また、以下に示す数値は既にこのサイト上に市区町村一覧の表として整理されて、MAPとして掲載されています。

ここでは年齢調整の効果が分かりやすいケースとして、ある県内市区町村 2ヶ所 A と B の女性のデータを比較してみます。まず図 1 に、神奈川県と A と B における被保険者全体の 1 人あたり総点数そのままの数値を示しました。県と比べて、A は約 6% 高く、B は約 5% 低いです。

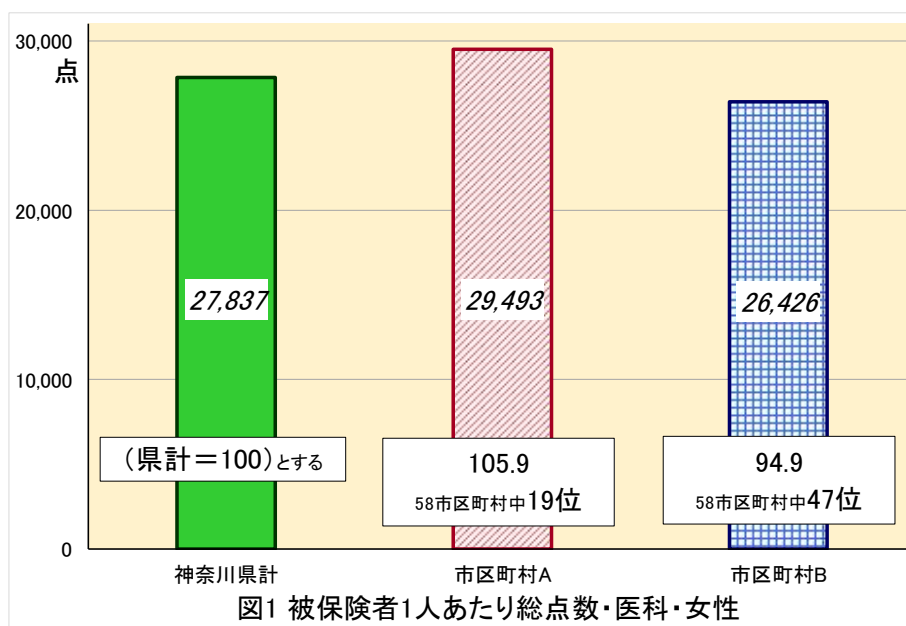
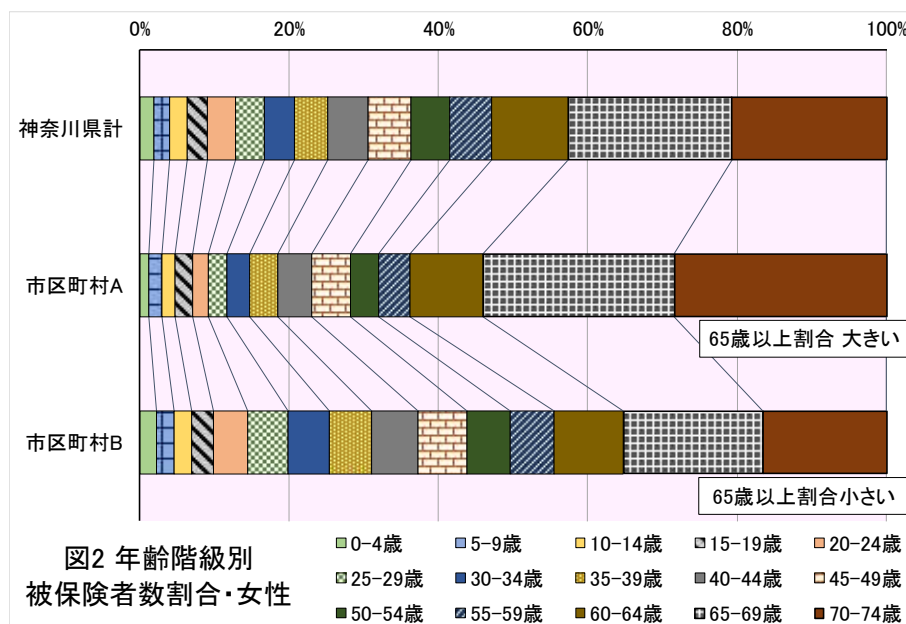
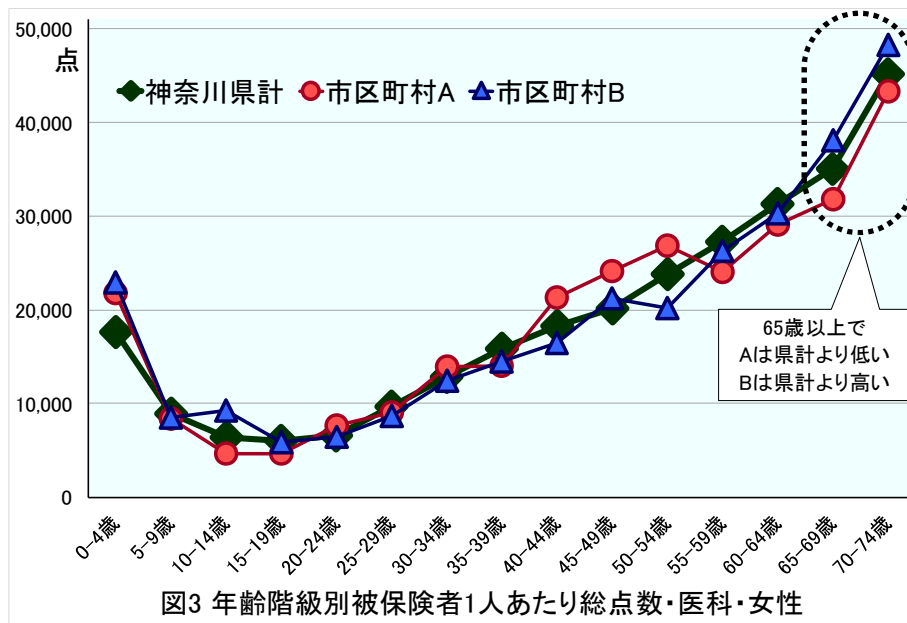


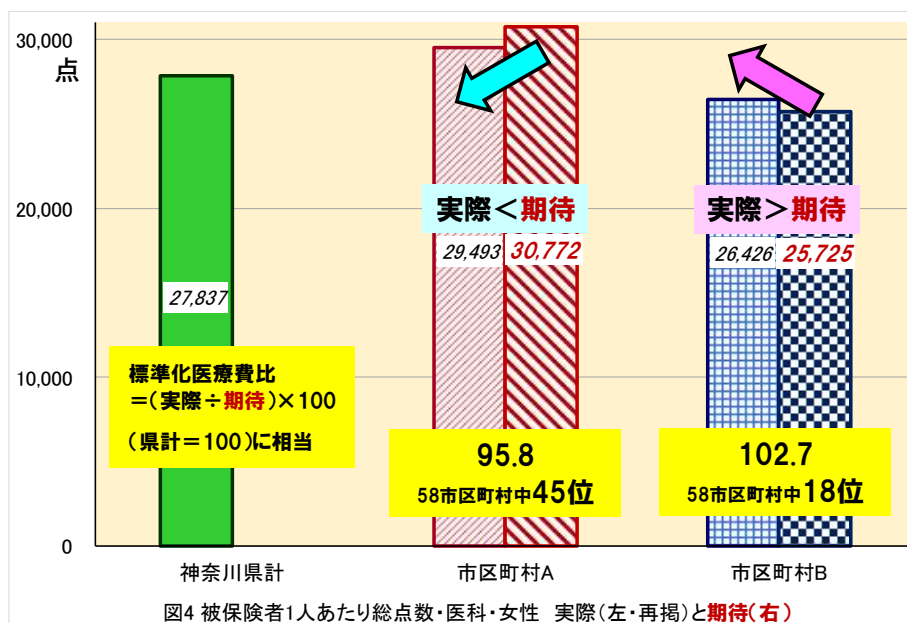
図 2 は、神奈川県と A と B の国保被保険者の年齢構成(5 歳階級別)割合です。県全体において、65 歳以上の割合は 4 割を超えます。県と比べて、被保険者のうち 65 歳以上の割合が A では大きく B では小さいです。



また、年齢(5 歳階級)階級別総点数のデータより、年齢階級ごとの被保険者 1 人あたり総点数を計算します(図 3)。年齢構成の主要な割合を占める 65 歳以上は、1 人あたり総点数も最も高く、結果として医療費にとって非常に決定的な集団です。なお今回の例では、県だけでなく A と B についても入手できました。



ここで、図3の神奈川県の年齢階級別被保険者1人あたり総点数を、図2の市区町村AあるいはBの国保被保険者の年齢構成にあてはめます(計算は文末の《補足》参照)。この意味は、「仮に、その市区町村の各年齢階級の被保険者1人あたり総点数が、ことごとく県全体と同じだったと想定」した時に、“期待”される被保険者全体の1人あたり総点数です。そして算出された「期待」の値を図1の「実際」と並べて示したものが図4です。Aの実際は期待より低く、Bの実際は期待より高いという結果になりました。



そして、「期待」と「実際」の比、具体的には「期待」の値を100とした時の「実際」の値の大きさが標準化医療費比です(簡単に「地域差指数」とも呼ばれます)。Aでは100より小さく、Bでは100より大きくなりました。繰り返しのようになりますが、そのままの数値はAで県よりも高く、Bで県よりも低いです。しかし、標準化医療費比によってAは年齢構成のわりに低く、Bは年齢構成のわりに高いと示されました。

ちなみに、標準化された数値は「対象集団と基準集団との比較」であり、対象集団間の比較は(年齢構成が全く異なる場合はとくに)お勧めできない、という厳密な考え方もあります。しかし、ここでの対象集団(各市区町村)は基準集団(神奈川県)の構成要素でもあることから、標準化された数値を比較の指標として使っています。

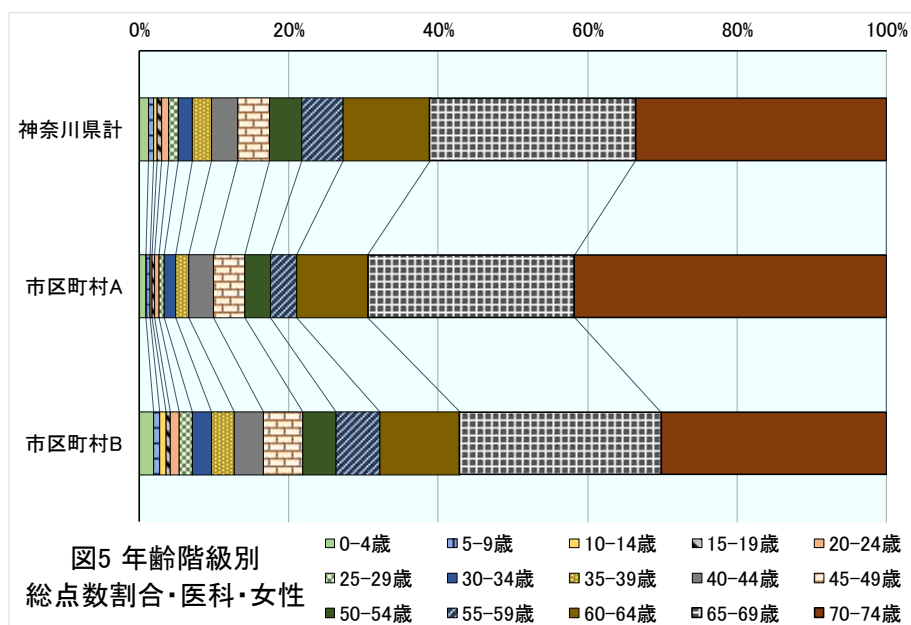
## 標準化計算の後に一可能ならば、年齢階級別の数値に立ちかえる

ここまで、被保険者 1 人あたり総点数を例にして年齢調整の手順を見てみました。そして、重要なのはこの先と言えます。標準化された数値によるひと目で分かる比較から、もし入手可能であれば年齢階級別のデータに立ち返ることで、差の「由来・出どころ」を明らかにしたいものです。

図 3 には、市区町村 A と B における年齢階級ごとの被保険者 1 人あたり総点数も示されています。65 歳以上において、A は県より低く B は県より高いことが分かります。つまり、A では 65 歳以上の被保険者が比較的多いが、その年齢層の 1 人あたり総点数は比較的低い。一方、B では 65 歳以上の被保険者が比較的小さいが、その 1 人あたり総点数は比較的高い。この違いが、標準化＝「年齢構成のわりには」での逆転をもたらしているようです。このように、標準化された数値を見た後で年齢階級別のデータに立ちかえるとよいでしょう。

今回の例は極めて単純な構図でした。ここでご紹介した手順は、ご存知の方にはきわめて当たり前のことかもしれません。

参考として、総点数の年齢(5 歳階級)別構成割合を示します(図 5)。県全体において、65 歳以上の構成割合は 6 割を超えています。また、図 3 で見られた 65 歳以上の人口の構成割合における A と B との間の差は、図 5 の総点数の構成割合においてはやや縮まっています。



## 《補足》標準化(間接法)の計算手順まとめ

県全体を基準とする場合に必要な元データは、以下の①～⑤の手順における「①県全体の年齢階級別のデータ(例の図3の県データに相当)」「②市(区)町村の年齢階級別人口(図2に相当)」「④市(区)町村の全体におけるデータ(図1に相当)」の3種類です。

算出は以下の手順で行われます。

- ① 県全体の年齢階級別のデータ(=人口に占める割合やひとりあたりの数値)を用意する。
- ② 市(区)町村の年齢階級別人口を用意する。
- ③ ②のような人口分布の集団(その市(区)町村)の状況が、仮に①(県全体)のデータと同じだったと仮定した時の、その集団全体(全年齢)における値=“期待XX”の値を算出する。

たとえば、上の例の場合、各年齢階級について順番に

$$(Aの0-4歳人口) \times (県全体の0-4歳1人あたり総点数) = (期待のAの0-4歳全員の総点数)$$

のように70-74歳まで計算します。次に、これらを足し算して

$$(期待のAの0-4歳全員の総点数) + (期待のAの5-9歳～) + \dots + (期待のAの70-74歳～) \\ = (期待のAの全員の総点数)$$

を計算します。今回の例の場合は、さらに、この数値を総人口で割り算して

$$(期待のAの全員の総点数) \div (Aの総人口) = (期待のA全体の1人あたり総点数)$$

を算出しています。

- ④ 市(区)町村の全体(全年齢)におけるデータを用意する。
- ⑤ ④と③の比を、③を100とした数値=“標準化XX比”として算出する。

③の数値の意味がややこしいかもしれません。文中の①②の位置の前後を入れ換えて全体を表現すれば、以下のようなになるでしょう。

県全体の状況(①)を、その市(区)町村の年齢構成(②)にあてはめた時に仮定(期待)される想定値=期待される値(③)を求める。その仮定の全体値に比べて、現実の市(区)町村の値(④)がどれだけ大きい/小さいか?を指標(⑤)とする。

市(区)町村の年齢階級別データ(例の図3のA・Bデータに相当)は、標準化の計算だけならば必要ありません。しかし入手できるならば、標準化XX比の値の理由・出どころを探るために立ちかえることが重要でしょう。