

平均寿命は生きた年数の平均値ではない

2024年7月30日

『そもそも論』の第24回は、「平均寿命は生きた年数の平均値ではない」です。

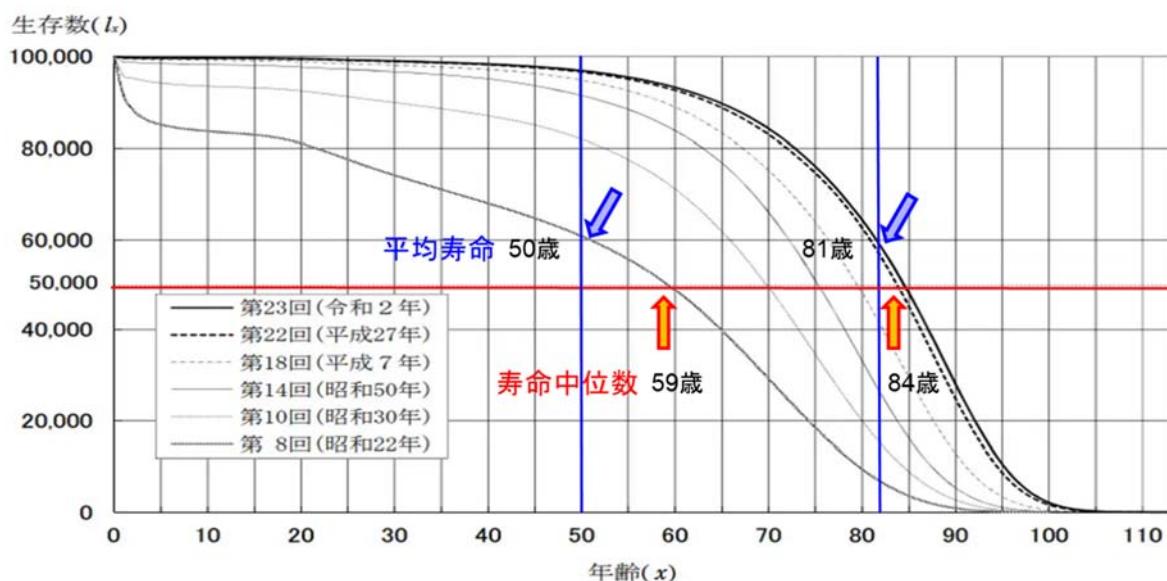
毎年公表される平均寿命。1947年には男性は50歳、女性は53歳でしたが[註1]、2023年には男性81歳、女性87歳に。ちなみに、明治・大正期は男女とも40歳代前半でした。「ずいぶん伸びたもんだ」と感嘆する方も多いでしょう。確かに寿命は伸びているのですが、本当に30年以上も長生きするようになったのでしょうか(というより、ひと昔前は本当に人生50年だったのでしょうか)。

[註1]1947年の平均寿命も1946年の人口動態データのみで算定されているので、出征や空襲など戦争の直接の影響は除かれています(ただし、戦争による傷病の後遺症や遅延発症、食糧不足等の影響は排除できません)。

寿命は生命表から算出するのですが、生命表は「前年の年齢ごとの死亡率を全年齢分を積み上げたもの」で、生きてきた過去の各年の該当年齢の死亡率を一生分並べたものではありません。つまり、「平均寿命」は、「昨年生まれたばかりの子が、昨年の各年齢の死亡率でこのまま生きていくと、どれくらい生きられそうか」という期待値であって、全ゼロ歳児の余命の平均値です。

もう一つの指標に「寿命中位数」があります。昨年一年間に生まれた子の「ちょうど半数が死亡するであろう年齢」を意味します(したがって全ゼロ歳児の余命の「中央値」と表現してもよいかもしれません)。こちらの方が感覚に近いでしょうね。ただし死亡確率を積み上げた期待値であることに変わりはありません。

図を見てください(死亡パターンの男女差は大きないので、ここでは一方のみを掲載します)。1947年(昭和22年)のように生まれた直後に多数が死亡する(乳幼児死亡が多い)と失われた年数がとても大きくなつて平均寿命が大きく引き下げられ(50歳)、半数が亡くなる寿命中位数



(59歳)と間に9年あまりの開きが出ます。乳幼児死亡が大きく減少した今日では平均寿命(81歳程度)と寿命中位数(84歳程度)の差は3年弱に縮まっています。

多数の数量データを要約するとき「代表値とばらつき」で示しますが、これを「平均値と標準偏差」で表示する場合はデータが正規分布することが前提になります。正規分布しないときは、対数で表すなどの関数変換をして正規化を図るか、それが不可能な場合は「中央値と四分位範囲(註2)」で表示します。寿命は正規分布が期待できないので、平均値で示すことは適切ではないように思います。

[註2]25パーセンタイル値と75パーセンタイル値(中央値は50パーセンタイル値とも言えます)

また、知りたいのは期待値(余命)ではなく「現に何歳ぐらいで人は亡くなっているか」という現状です。それを表す指標に、今年亡くなった人の年齢の平均値や中央値と、今年亡くなった人の年齢の最頻値(死亡数最大年齢or死亡ピーク年齢)があります。これらは、年齢別人口の影響を受け、出生数が多いいわゆる“団塊の世代”(1947～1949年生まれ)や“団塊ジュニア世代”(1971～1974年生まれ)の影響が強く出ますし、反対に“丙午(ひのえうま)年”(1906年、1966年)生まれは出生者が少なくて影響が小さくなりますので、これらもまた(実測値ではあるけれど)寿命の年次推移を見るのに最適というわけではありません。

ということで、たかが寿命を知りたいだけなのですが、ずいぶん苦労するのです。