# 平均寿命の計算について

みなさんは、こんなニュースを一度くらいは聞いたことがあるかと思います。

日本経済新聞・平成 14 年 8 月 1 日の記事

日本人の平均寿命は男性が 78.07 歳、女性は 84.93 歳となり、過去最高を更新したことが 31 日、厚生労働省がまとめた「平成 13 年簡易生命表」で分かった。男女とも世界一となる見込み。平成 13 年に生まれた赤ちゃんのうち、男性の半数以上、女性の四人に三人が八十歳の誕生日を迎える計算になり、今世紀後半の長寿社会が浮かび上がった。

ここで登場した平均寿命という言葉に対して、私は学生時代に勘違いをしていました。平均寿命とは、ここ 1 年間に亡くなった方の死亡時の年齢の平均だと思っていたのです。しかし、私の仕事との関係で勉強を進めるうちに、これは勘違いであることに気付きました。実際の平均寿命の定義を言葉で書くと、次のようになります。

平均寿命の定義

生命表における平均寿命とは、現在における死亡状況が今後変化しないと仮定したときに、今後出生する人が何年生きられるかという期待値である。

本稿では、平均寿命の計算を通して、「生命表」というものを紹介していきます。

以下、${}_tq_x$, ${}_tp_x$, $l_x$, $d_x$, $\mathring{e}_x$ という「生命関数」を定義します。厚生労働省や生命保険会社の文献では、他にも、$L_x$, $T_x$ などの記号があるのですが、ここでは省略します。

**死亡率** ${}_tq_x$　ちょうど $x$ 才の人が $t$ 年以内に死亡する確率。$t=1$ の場合、${}_1q_x = q_x$ と書く。

**生存率** ${}_tp_x$　ちょうど $x$ 才の人が $t$ 年後に生存している確率。${}_tp_x = 1 - {}_tq_x$, ${}_{s+t}p_x = {}_sp_{x+t} \cdot {}_tp_x$, ${}_0p_x = 1$ が成り立つ。$t=1$ の場合、${}_1p_x = p_x$ と書き、$p_x = 1 - q_x$ が成り立つ。

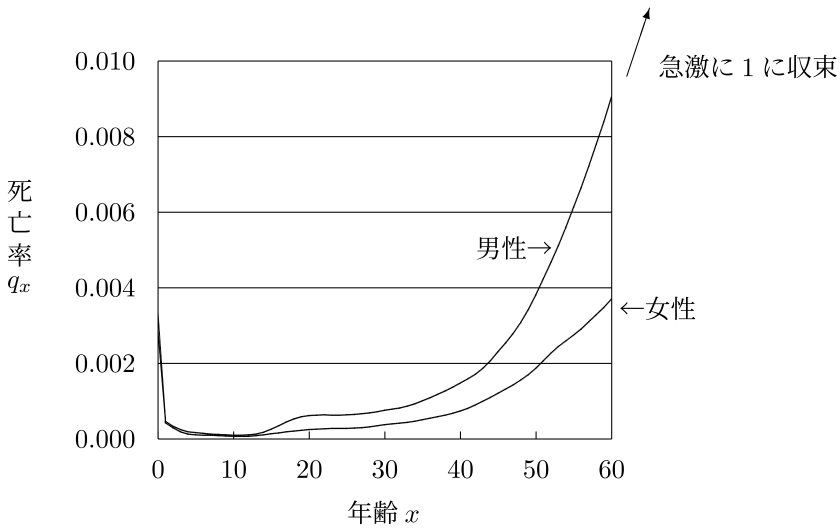
**生存数** $l_x$　今、10 万人の人が同時に生まれ、上記の $q_x$ に従って人口が減少していくとした場合、$x$ 年後に生存している人の数。$l_0 = 100000$, $l_{x+t} = l_x \ {}_tp_x$ が成り立つ。

**死亡数** $d_x$　$x$ 才における生存数 $l_x$ 人のうち、1 年以内に死亡する人の数。$d_x = l_xq_x = l_x - l_{x+1}$ が成り立つ。

実際の統計データとしては、まず、非負の整数 $x$ に対して、$\{q_x\}_{x=0,1,2,\cdots}$ が与えられています。このデータには、過去 1 年間の観察期間で、$x$ 才の人のうち亡くなってしまった方の割合を使います。分母、分子をどう決めるかとか、「ちょうど」$x$ 才の人の死亡率をどう計算するかは、ここでは省略させてください。さて、$\{q_x\}$ が決まれば、$p_x = 1 - q_x$, $l_0 = 100000$, $l_{x+1} = l_xp_x$, $d_x = l_xq_x$ によって、$\{p_x\}$,$\{l_x\}$,$\{d_x\}_{x=0,1,2,\cdots}$ が順に決まっていきます。平成 13 年簡易生命表の男性の場合は、この結果が次のような表になります。$\mathring{e}_x$ という項目もありますが、これが「平均余命」であり、以下で説明していきます。

平成 13 年簡易生命表（男性）より抜粋

| $x$ | $q_x$ | $p_x$ | $l_x$ | $d_x$ | $\mathring{e}_x$ |
|---|---|---|---|---|---|
| 0 | 0.00330 | 0.99670 | 100000 | 330 | 78.07 |
| | ⋮ | | ⋮ | | |
| 20 | 0.00062 | 0.99938 | 99201 | 62 | 58.64 |
| | ⋮ | | ⋮ | | |
| 80 | 0.06069 | 0.93931 | 53520 | 3248 | 8.13 |
| 81 | 0.06793 | 0.93207 | 50272 | 3415 | 7.62 |
| 82 | 0.07617 | 0.92383 | 46857 | 3569 | 7.14 |
| | ⋮ | | ⋮ | | |
| 100 | 1.00000 | 0.00000 | 1206 | 1206 | 2.20 |




生命表を見てすぐにわかるのは、死亡率は乳幼児のうちは高く、10 才くらいまで単調減少し、それ以降は単調増加して、高齢になって 1 に収束していくという点です。また、20 代男性の死亡率が比較的高くなっているのは、「不慮の事故」や「自殺」による死亡が多いためと言われています。なお、冒頭の新聞記事の後半は、男性については $\dfrac{l_{80}}{l_0} > \dfrac{1}{2}$、女性については $\dfrac{l_{80}}{l_0} > \dfrac{3}{4}$ を意味しています。

これらの記号を用いて、平均余命 $\mathring{e}_x$ と平均寿命を定義できます。

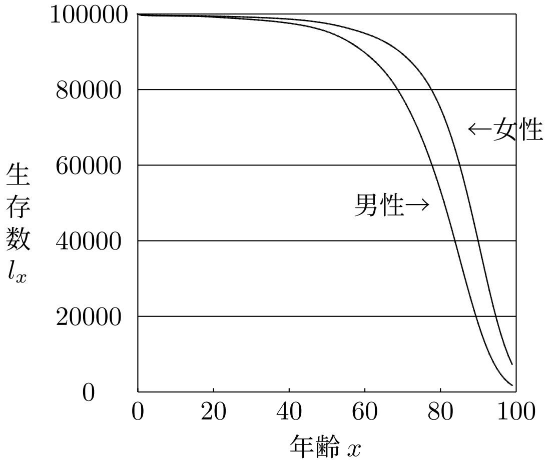
**平均余命** $\mathring{e}_x$　各年齢の死亡率 $q_x$ が今後一定であるとした場合の $x$ 才の人の亡くなるまでの年数（連続値）の期待値。

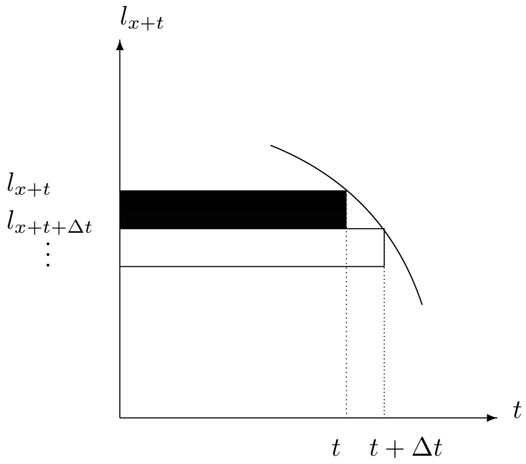
**平均寿命**　生まれたばかりの 0 才児の平均余命 $\mathring{e}_0$ のこと。

実は $\mathring{e}_x$ は前述の生命関数の積分を用いて書くことができます。しかしその前に注意しておきたいのが、$l_x$ はいま、非負の整数 $x$ に対して定義されていましたが、数学的な補間を用いて $x$ の連続関数に拡張して考えるという点です。補間には、4 次関数や指数関数を組み合わせた方法が採られています。これにより、$t, x$ の関数である ${}_tp_x = \dfrac{l_{x+t}}{l_x}$ も連続な関数に拡張されることになります。

さて、いま、$x$ 才の人が $l_x$ 人いると仮定すると、このうち経過 $t$ 年と $t+\Delta t$ 年 $(\Delta t \ll 1)$ の間に亡くなる方は $l_{x+t} - l_{x+t+\Delta t}$ 人で、この人たちの余命は $x$ 才時点で $t$ $(\approx t+\Delta t)$ だったことになり

ます。よって、この人たちの死亡時の年齢の和は、$t(l_{x+t} - l_{x+t+\Delta t})$ で、右下の図の黒い部分の面積にあたります。




同じことを $\Delta t$ 間隔ですべての $t$ について 0 から $\infty$ まで行って足し合わせると、$l_x$ 人全員の死亡時の年齢の総和が求まるわけです。$\Delta t \to 0$ とすると、実はこれが $\int_0^\infty l_{x+t}dt$ になっていることは図より明らかかと思います。図を使わないのであれば、部分積分を用いて、

$$\begin{aligned}\sum_{\substack{t=0\\ \Delta t\ \text{間隔}}}^{\infty} t(l_{x+t} - l_{x+t+\Delta t}) &\approx \int_0^\infty t\left(-\frac{d}{dt}l_{x+t}\right)dt \\ &= [t(-l_{x+t})]_0^\infty - \int_0^\infty (-l_{x+t})dt \\ &= \int_0^\infty l_{x+t}dt\end{aligned}$$

と書けば少しは厳密でしょうか。いずれにしても、これを $l_x$ で割れば $t=0$ での余命の平均が出ますので、次が求まりました。

平均余命の積分による表示

$$\mathring{e}_x = \frac{1}{l_x}\int_0^\infty l_{x+t}dt = \int_0^\infty {}_tp_x dt$$

実際、平均寿命はこの式により計算されていることが、簡易生命表の前書きにも載っています。直感の優れた人なら、この式だけで「確かに余命の期待値だな」と思えるかもしれませんね。

ところで、積分を台形近似し、${}_0p_x = 1$ にも注意すると、

$$\mathring{e}_x \approx \sum_{t=0}^{\infty} \frac{{}_tp_x + {}_{t+1}p_x}{2} = \frac{1}{2} + \sum_{t=1}^{\infty} {}_tp_x$$

が得られます。さらに、この近似のもとでは、

$$\begin{aligned} \mathring{e}_x &\approx \frac{1}{2} + p_x \sum_{t=1}^{\infty} {}_{t-1}p_{x+1} \\ &= \frac{1}{2} + p_x\left(1+\sum_{t=1}^{\infty} {}_{t}p_{x+1}\right) \\ &= \frac{1}{2} + p_x\left(1+\mathring{e}_{x+1} - \frac{1}{2}\right) \\ &= \left(1+\mathring{e}_{x+1}\right)p_x + \frac{1}{2}q_x \end{aligned}$$

という（逆向きの）漸化式が得られます。この式は、$x$ 才の人が 1 年以内に死なない場合の平均余命 $1+\mathring{e}_{x+1}$ と、1 年以内に死ぬ場合の余命の近似値 $\frac{1}{2}$ を用いた条件付期待値の式になっています。この式があれば、積分を使わなくても、$q_x$ から $\mathring{e}_x$ の近似が求まります。

上の近似式を用いて、平成 13 年簡易生命表の死亡率 $\{q_x\}_{x=0,1,2,\cdots}$ から、平均寿命 $\mathring{e}_0$ を推計してみましょう。具体的には、$q_{100}=1, p_{100}=0$ より、最初だけ $\mathring{e}_{100}=0.5$ と仮定してしまい、以下は漸化式により、$q_{99}, p_{99}, \mathring{e}_{100}$ から $\mathring{e}_{99}$ を算出し、次は $q_{98}, p_{98}, \mathring{e}_{99}$ から $\mathring{e}_{98}$ を算出し、と繰り返し $\mathring{e}_{99}, \mathring{e}_{98}, \cdots$ を算出していきます。すると最後に、$\mathring{e}_0=78.05$ が得られます。台形近似に依らない厚生労働省の公表値が 78.07 ですので、ほぼ再現できているといえます。近似の誤差の原因の一つとしては、平成 13 年簡易生命表の高齢部分が考えられます。厚生労働省は、100 歳以上の部分の詳細な $l_x$ を用いて $\mathring{e}_{100}=2.2$ を算出しているにも関わらず、その後に改めて、$q_{100}=1, l_{100}=d_{100}$ と置いたものを発表しているからです。5 年に 1 回発行される「完全生命表」でならこんなことは起こらず、私もたまに仕事で近似を使います。

平均寿命の計算ぐらいで、厚生労働省が積分を使っているなんて、知る人は少なかったんではないでしょうか。しかも、探れば探るほど、泥臭い計算が緻密に行われていることがわかってきます。ふとしたニュースでも、みなさんに技術的な側面、確率的な側面を提示できれば幸いです。

平均余命の積分表示からも明らかですが、平均寿命、平均余命が延びるというのは、つまり、ほぼ全年齢において死亡率が低下することを意味しています。生命保険会社においては、国民死亡率を観察していくことが非常に重要です。安直に、「死亡率が下がる方が、保険会社にとっては保険金支払いが減って得になるんじゃないの？」と思われる方もいるかもしれません。確かに死亡保険についてはそうなんですが、生命保険会社では個人年金や医療保険という商品も扱っており、こちらにとっては死亡率の低下は脅威です。例えば終身年金という、「お客様が生存している限り、毎年一定額を支払います」という契約については、長生きが保険会社にとってリスクになってしまいます。生命保険会社としては、個々の商品について収支を安定させるように努めます。そのため、国民死亡率に大きな変化があれば、新規の契約に対して、保険料の計算基礎となる死亡率を変更していきます。

この文章のの参考文献は、「平成 13 年簡易生命表」（厚生労働省大臣官房統計情報部編）です。下記の URL から EXCEL ファイルをダウンロードすることもできます。

http://www.mhlw.go.jp/toukei/saikin/hw/life/life01/index.html

成川淳 (なるかわあつし)