Ⅱ　令和２年市区町村別生命表の作成方法

１．対象

各市区町村における日本人とした。

２．作成基礎期間

平成 31 年 1 月 1 日から令和３年 12 月 31 日に至る 3 年間とした。

３．基礎資料

市区町村別に次の資料を用いた。なお、市区町村の名称及び区域は、令和３年 12 月 31 日現在のものとした。

(1)　令和元年～３年　性・年齢階級別死亡数
　　　　－厚生労働省政策統括官（統計・情報政策、労使関係担当）　人口動態統計－

(2)　令和２年７月～9 月　性・年齢階級別死亡数
　　　　－厚生労働省政策統括官（統計・情報政策、労使関係担当）　人口動態統計－

(3)　平成 30 年～令和３年　性別出生数
　　　　－厚生労働省政策統括官（統計・情報政策、労使関係担当）　人口動態統計－

(4)　令和２年 10 月 1 日現在　性・年齢別人口（不詳補完結果（参考表））
　　　　－総務省統計局　令和２年国勢調査－

実際の計算にあたっての、上記資料の利用区分及び後述の計算方法の説明に用いる記号を次に示す（以下、記号の混乱を避けるため、それが都道府県、指定都市及び東京都区部 i の数値を表す場合は右肩に i、i に属する市区町村 j の数値を表す場合は (i,j) と、それぞれ付して区別する）。

人口及び死亡数

年齢	人口（令和２年 10 月 1 日）		死亡数	
	年齢階級別の人口	各年齢階級の始年齢人口	令和２年7 月～9 月	平成 31 年（令和元年）～３年3 か年合計
0 歳	…	…	…	$D_0^{(i,j)}$
1～4	${}_4\overset{*}{P}{}_1^{(i,j)}$	$\overset{*}{P}{}_1^{(i,j)}$	${}_4\overset{*}{D}{}_1^{(i,j)}$	${}_4D_1^{(i,j)}$
5～9	${}_5\overset{*}{P}{}_5^{(i,j)}$	$\overset{*}{P}{}_5^{(i,j)}$	${}_5\overset{*}{D}{}_5^{(i,j)}$	${}_5D_5^{(i,j)}$

10～14	${}_5\overset{*}{P}{}_{10}^{(i,j)}$	$\overset{*}{P}{}_{10}^{(i,j)}$	${}_5\overset{*}{D}{}_{10}^{(i,j)}$	${}_5D_{10}^{(i,j)}$
⋮	⋮	⋮	⋮	⋮
x～$x+4$	${}_5\overset{*}{P}{}_{x}^{(i,j)}$	$\overset{*}{P}{}_{x}^{(i,j)}$	${}_5\overset{*}{D}{}_{x}^{(i,j)}$	${}_5D_{x}^{(i,j)}$
⋮	⋮	⋮	⋮	⋮
90～94	${}_5\overset{*}{P}{}_{90}^{(i,j)}$	$\overset{*}{P}{}_{90}^{(i,j)}$	${}_5\overset{*}{D}{}_{90}^{(i,j)}$	${}_5D_{90}^{(i,j)}$
95～99		$\overset{*}{P}{}_{95}^{(i,j)}$	${}_5\overset{*}{D}{}_{95}^{(i,j)}$	
	…			…

（注意）この基礎資料には年齢や住所地が不詳であるデータが存在しているので、次のように不詳按分をすることによりデータの補正を行った。４．で述べる計算はこの補正されたデータに対して行った。

① **人口の補正**

不詳補完結果（参考表）を用いた。

② **死亡数の補正**

死亡数には、「(a) 年齢のみ不詳、(b) 住所地のみ不詳、(c) 年齢及び住所地ともに不詳」という 3 つの不詳のパターンが存在する。これらの不詳按分は次のような順で行った。

(a)　年齢のみ不詳

年齢及び住所地ともに既知の死亡数をもとに、市区町村ごとに年齢別死亡数に比例させて按分して加えた。

(b)　住所地のみ不詳

(a) の結果をもとに、年齢ごとに市区町村別死亡数に比例させて按分して加えた。

(c)　年齢及び住所地ともに不詳

(b) の結果をもとに、年齢・市区町村別死亡数に比例させて按分して加えた。

出生数

平成 30 年～令和２年	平成 31 年（令和元年）～３年
$B^{(i,j)}\binom{'18}{'20}$	$B^{(i,j)}\binom{'19}{'21}$

４．計算方法

今回の市区町村別生命表の作成においては、Chin Long Chiang の方法によった。この方法は、各国で比較的多く用いられている作成方法である。

（１）死亡率推定のための地域

都道府県、指定都市及び東京都区部とした。

（２）中央人口 ${}_nP_x^{(i,j)}$

作成基礎期間の中央にあたる令和２年 7 月 1 日現在の人口を中央人口という。市区町村 (i,j) の 1 歳以上 5 歳未満の中央人口を ${}_4P_1^{(i,j)}$、x 歳以上 $x+5$ 歳未満の中央人口を ${}_5P_x^{(i,j)}$（ただし $x=5,10,15,\cdots,90$）で表すとき、それらを、

$$
{}_4P_1^{(i,j)} = {}_4\overset{*}{P}{}_1^{(i,j)} + \frac{1}{4}\left(\overset{*}{P}{}_5^{(i,j)} - \overset{*}{P}{}_1^{(i,j)}\right) + \frac{31}{32}\,{}_4\overset{*}{D}{}_1^{(i,j)} + \frac{1}{40}\,{}_5\overset{*}{D}{}_5^{(i,j)}
$$

$$
{}_5P_x^{(i,j)} = {}_5\overset{*}{P}{}_x^{(i,j)} + \frac{1}{4}\left(\overset{*}{P}{}_{x+5}^{(i,j)} - \overset{*}{P}{}_x^{(i,j)}\right) + \frac{39}{40}\,{}_5\overset{*}{D}{}_x^{(i,j)} + \frac{1}{40}\,{}_5\overset{*}{D}{}_{x+5}^{(i,j)}
$$

（$x=5,10,15,\cdots,90$）

により求めた。

（３）0 歳の死亡率 $q_0^{(i,j)}$

市区町村 (i,j) ごとに、0 歳の粗死亡率 $\tilde{q}_0^{(i,j)}$ を、

$$
\tilde{q}_0^{(i,j)} = \frac{D_0^{(i,j)}}{\frac{1}{2}\left\{B^{(i,j)}\binom{'18}{'20} + B^{(i,j)}\binom{'19}{'21}\right\}}
$$

により求めた。

次に、地域 i ごとに、

$$
D_0^i = \sum_j D_0^{(i,j)}、B^i\binom{'18}{'20} = \sum_j B^{(i,j)}\binom{'18}{'20}、B^i\binom{'19}{'21} = \sum_j B^{(i,j)}\binom{'19}{'21}
$$

とし、粗死亡率の人口による重み付き平均 E_0^i 及び分散 V_0^i を、

$$
E_0^i = \sum_j \tilde{q}_0^{(i,j)} \times \frac{\frac{1}{2}\left\{B^{(i,j)}\binom{'18}{'20} + B^{(i,j)}\binom{'19}{'21}\right\}}{\frac{1}{2}\left\{B^i\binom{'18}{'20} + B^i\binom{'19}{'21}\right\}} = \frac{D_0^i}{\frac{1}{2}\left\{B^i\binom{'18}{'20} + B^i\binom{'19}{'21}\right\}}
$$

$$
V_0^i = \sum_j \left(\tilde{q}_0^{(i,j)}\right)^2 \times \frac{\frac{1}{2}\left\{B^{(i,j)}\binom{'18}{'20} + B^{(i,j)}\binom{'19}{'21}\right\}}{\frac{1}{2}\left\{B^i\binom{'18}{'20} + B^i\binom{'19}{'21}\right\}} - \left(E_0^i\right)^2
$$

で求め、パラメータ α_0^i、β_0^i を、

$$
\alpha_0^i = E_0^i\left\{\frac{E_0^i\left(1-E_0^i\right)}{V_0^i} - 1\right\}、\beta_0^i = \left(1-E_0^i\right)\left\{\frac{E_0^i\left(1-E_0^i\right)}{V_0^i} - 1\right\}
$$

とおいた。

最後に、市区町村 (i,j) ごとに、地域 i のパラメータ α_0^i、β_0^i を用いて、５．で述べるベイズ推定の考え方から、死亡率 $q_0^{(i,j)}$ を、

$$q_0^{(i,j)} = \frac{\alpha_0^i + D_0^{(i,j)}}{\alpha_0^i + \beta_0^i + \frac{1}{2}\left\{B^{(i,j)}\binom{'18}{'20} + B^{(i,j)}\binom{'19}{'21}\right\}}$$

により求めた。

（４）1 歳以上の死亡率 ${}_nq_x^{(i,j)}$

市区町村 (i,j) ごとに、1 歳以上の年齢階級ごとの粗中央死亡率 ${}_n\widetilde{m}_x^{(i,j)}$ を、

$${}_n\widetilde{m}_x^{(i,j)} = \frac{{}_nD_x^{(i,j)}}{3\cdot{}_nP_x^{(i,j)}}$$

により求めた。ただし、年齢階級は $x = 1,5,10,15,\cdots,90$ として、それぞれ $x = 1$ のとき $n = 4$、$x = 5,10,15,\cdots,90$ のとき $n = 5$とした（以下同様）。

次に、地域 i ごとに、

$${}_nD_x^i = \sum_j {}_nD_x^{(i,j)}、\ {}_nP_x^i = \sum_j {}_nP_x^{(i,j)}$$

とし、粗中央死亡率の人口による重み付き平均 ${}_nE_x^i$ 及び分散 ${}_nV_x^i$ を、

$${}_nE_x^i = \sum_j {}_n\widetilde{m}_x^{(i,j)} \times \frac{{}_nP_x^{(i,j)}}{{}_nP_x^i} = \frac{{}_nD_x^i}{3\cdot{}_nP_x^i}、\ {}_nV_x^i = \sum_j \left({}_n\widetilde{m}_x^{(i,j)}\right)^2 \times \frac{{}_nP_x^{(i,j)}}{{}_nP_x^i} - \left({}_nE_x^i\right)^2$$

で求め、パラメータ ${}_n\alpha_x^i$、${}_n\beta_x^i$ を、

$${}_n\alpha_x^i = {}_nE_x^i\left\{\frac{{}_nE_x^i\left(1 - {}_nE_x^i\right)}{{}_nV_x^i} - 1\right\}、\ {}_n\beta_x^i = \left(1 - {}_nE_x^i\right)\left\{\frac{{}_nE_x^i\left(1 - {}_nE_x^i\right)}{{}_nV_x^i} - 1\right\}$$

とおいた。

最後に、市区町村 (i,j) ごとに、地域 i のパラメータ ${}_n\alpha_x^i$、${}_n\beta_x^i$ を用いて、５．で述べるベイズ推定の考え方から、中央死亡率 ${}_nm_x^{(i,j)}$ を、

$${}_nm_x^{(i,j)} = \frac{{}_n\alpha_x^i + {}_nD_x^{(i,j)}}{{}_n\alpha_x^i + {}_n\beta_x^i + 3\cdot{}_nP_x^{(i,j)}} \quad (x = 1,5,10,15,\cdots,90)$$

により求めた。

こうして得られた中央死亡率 ${}_nm_x^{(i,j)}$ を、変換式

$${}_nq_x^{(i,j)} = \frac{n\cdot{}_nm_x^{(i,j)}}{1 + \left(n - {}_na_x^i\right){}_nm_x^{(i,j)}} \quad (x = 1,5,10,15,\cdots,90)$$

で変換することにより、死亡率 ${}_nq_x^{(i,j)}$ を求めた。

ただしここで、${}_na_x^i$ は、地域 i の平均生存期間であり、令和２年都道府県別生命表における生存数 l_{x+n}^i、死亡数 ${}_nd_x^i$ 及び定常人口 ${}_nL_x^i$ から、

$${}_na_x^i = \frac{{}_nL_x^i - n\cdot l_{x+n}^i}{{}_nd_x^i}$$

により定義されるものである（$x=0$ のときは $n=1$ 、$x=95$のときは$n=\infty$とする（以下同様））。

（５）生存数 $l_x^{(i,j)}$ 及び死亡数 ${}_nd_x^{(i,j)}$

0 歳の生存数を $l_0^{(i,j)}=100{,}000$ とし、生存率を逐次乗じることで、

$$
\begin{array}{lll}
p_0^{(i,j)}=1-q_0^{(i,j)} & l_1^{(i,j)}=l_0^{(i,j)}\times p_0^{(i,j)} & d_0^{(i,j)}=l_0^{(i,j)}-l_1^{(i,j)} \\
{}_4p_1^{(i,j)}=1-{}_4q_1^{(i,j)} & l_5^{(i,j)}=l_1^{(i,j)}\times {}_4p_1^{(i,j)} & {}_4d_1^{(i,j)}=l_1^{(i,j)}-l_5^{(i,j)} \\
{}_5p_5^{(i,j)}=1-{}_5q_5^{(i,j)} & l_{10}^{(i,j)}=l_5^{(i,j)}\times {}_5p_5^{(i,j)} & {}_5d_5^{(i,j)}=l_5^{(i,j)}-l_{10}^{(i,j)} \\
{}_5p_{10}^{(i,j)}=1-{}_5q_{10}^{(i,j)} & l_{15}^{(i,j)}=l_{10}^{(i,j)}\times {}_5p_{10}^{(i,j)} & {}_5d_{10}^{(i,j)}=l_{10}^{(i,j)}-l_{15}^{(i,j)} \\
\cdots & \cdots & \cdots \\
{}_5p_{90}^{(i,j)}=1-{}_5q_{90}^{(i,j)} & l_{95}^{(i,j)}=l_{90}^{(i,j)}\times {}_5p_{90}^{(i,j)} & {}_5d_{90}^{(i,j)}=l_{90}^{(i,j)}-l_{95}^{(i,j)} \\
 & & {}_\infty d_{95}^{(i,j)}=l_{95}^{(i,j)}
\end{array}
$$

により生存数 $l_x^{(i,j)}$ 及び死亡数 ${}_nd_x^{(i,j)}$ を求めた。

（６）定常人口 ${}_nL_x^{(i,j)}$

x 歳以上 $x+n$ 歳未満の定常人口 ${}_nL_x^{(i,j)}$ は、

$$
{}_nL_x^{(i,j)}=n\cdot l_{x+n}^{(i,j)}+{}_na_x^i\cdot{}_nd_x^{(i,j)}(x=0,1,5,10,15,\cdots,90)
$$
$$
{}_\infty L_{95}^{(i,j)}={}_\infty a_{95}^i\cdot{}_\infty d_{95}^{(i,j)}
$$

により求めた。

（７）定常人口 $T_x^{(i,j)}$

x 歳以上の定常人口 $T_x^{(i,j)}$ は、

$$
T_x^{(i,j)}=\sum_{t=x}^{95}{}_nL_t^{(i,j)}
$$

により求めた。

（８）平均余命 $\mathring{e}_x^{(i,j)}$

平均余命 $\mathring{e}_x^{(i,j)}$ は、

$$
\mathring{e}_x^{(i,j)}=\frac{T_x^{(i,j)}}{l_x^{(i,j)}}
$$

により求めた。

（９）平均余命 $\mathring{e}_x^{(i,j)}$ の標準誤差

平均余命 $\mathring{e}_x^{(i,j)}$ の標準誤差を、中央死亡率（0 歳では死亡率）の事後分布の分散（５．を参照）

$$ {}_{n}V_x^{(i,j)} = \frac{\left({}_{n}\alpha_x^i + {}_{n}D_x^{(i,j)}\right)\left({}_{n}\beta_x^i + 3\cdot {}_{n}P_x^{(i,j)} - {}_{n}D_x^{(i,j)}\right)}{\left({}_{n}\alpha_x^i + {}_{n}\beta_x^i + 3\cdot {}_{n}P_x^{(i,j)}\right)^2\left({}_{n}\alpha_x^i + {}_{n}\beta_x^i + 3\cdot {}_{n}P_x^{(i,j)} + 1\right)} $$

（ただし $x=0$ では $3\cdot P_0^{(i,j)} = \frac{1}{2}\left\{B^{(i,j)}\binom{'18}{'20} + B^{(i,j)}\binom{'19}{'21}\right\}$ とする）

を用いて、

$$ \frac{1}{l_x^{(i,j)}}\sqrt{\sum_{t=x}^{90}\left(l_t^{(i,j)}\right)^2\cdot\left(n - {}_{n}a_t^i + \mathring{e}_{t+n}^{(i,j)}\right)^2\cdot {}_{n}V_t^{(i,j)}} $$

$(x = 0,1,5,10,15,\cdots,90)$

により求めた。

５．死亡率のベイズ推定

生命表作成の基礎となる中央死亡率（0 歳では死亡率）については、市区町村ごとでは死亡数が少なく、人口（0 歳では出生数）と死亡数の実績値から直接計算しただけでは偶然変動の影響を受けて不安定であることが多いため、ベイズ統計学の手法を用いた推定（ベイズ推定）を行った。

ベイズ推定とは、前もって利用可能な情報を事前分布として表現したものに、観測によって得られる標本情報を尤度として乗ずることで、ベイズの定理により決定される事後分布に基づき推定する方法をいう。

（１）ベイズ推定の考え方

母数 θ（ここでは市区町村 (i,j) の死亡率を指す）の取り得る値について、あらかじめ何らかの情報（ここでは地域 i 内の死亡率の分布を指す）が与えられているとき、それを確率分布の形で表現したものを事前分布という。

θ の事前分布の確率密度関数を $p(\theta)$ とし、また、θ を与えたときのデータ X（ここでは市区町村 (i,j) の死亡数を指す）の確率密度関数を $f(X|\theta)$ とすると、データの観測値 $X = x$ が与えられたときの θ の条件付き確率密度関数 $p(\theta|x)$ は、ベイズの定理より、

$$ p(\theta|x) = \frac{p(\theta)\cdot f(x|\theta)}{\int p(\theta)\cdot f(x|\theta)d\theta} $$

で与えられる。これを θ の事後確率密度関数、その分布を事後分布という。

ここで分母の正規化定数を無視すれば、$p(\theta|x) \propto p(\theta)\cdot f(x|\theta)$ となっており、事前分布 $p(\theta)$ にデータの観測値の情報を θ の尤度関数 $f(x|\theta)$ として乗ずることにより、事後分布 $p(\theta|x)$ が決定されることが分かる。

（２）死亡率の推定方法

具体的には、以下に述べる方法で死亡率を推定した。

1 歳以上の死亡率については、母数を中央死亡率 ${}_{n}m_x^{(i,j)}$ とし、その事前分布として、次

の性質を持つベータ分布 $Beta\left({}_n\alpha_x^i, {}_n\beta_x^i\right)$ を選択した。すなわち、

$$p\left({}_nm_x^{(i,j)}\right) = \frac{1}{B\left({}_n\alpha_x^i, {}_n\beta_x^i\right)}\left({}_nm_x^{(i,j)}\right)^{{}_n\alpha_x^i-1}\left(1-{}_nm_x^{(i,j)}\right)^{{}_n\beta_x^i-1}$$

とすると、その平均及び分散は $\dfrac{{}_n\alpha_x^i}{{}_n\alpha_x^i + {}_n\beta_x^i}$ 及び $\dfrac{{}_n\alpha_x^i \cdot {}_n\beta_x^i}{\left({}_n\alpha_x^i + {}_n\beta_x^i\right)^2\left({}_n\alpha_x^i + {}_n\beta_x^i + 1\right)}$ で与えられるが、これらが粗中央死亡率の人口による重み付き平均 ${}_nE_x^i$ 及び分散 ${}_nV_x^i$ に等しくなるよう、事前分布のパラメータを、それぞれ、

$${}_n\alpha_x^i = {}_nE_x^i\left\{\frac{{}_nE_x^i\left(1-{}_nE_x^i\right)}{{}_nV_x^i} - 1\right\}、{}_n\beta_x^i = \left(1-{}_nE_x^i\right)\left\{\frac{{}_nE_x^i\left(1-{}_nE_x^i\right)}{{}_nV_x^i} - 1\right\}$$

により決定した。

ここで中央人口 ${}_nP_x^{(i,j)}$ は既知とし、死亡数 ${}_nD_x^{(i,j)}$ は２項分布 $Bin\left(3\cdot{}_nP_x^{(i,j)}, {}_nm_x^{(i,j)}\right)$ に従う確率変数 ${}_n\widetilde{D}_x^{(i,j)}$ の実現値と考える。このとき ${}_n\widetilde{D}_x^{(i,j)} = {}_nD_x^{(i,j)}$ となる確率は、

$$f\left({}_nD_x^{(i,j)}\middle|{}_nm_x^{(i,j)}\right) = \binom{3\cdot{}_nP_x^{(i,j)}}{{}_nD_x^{(i,j)}}\left({}_nm_x^{(i,j)}\right)^{{}_nD_x^{(i,j)}}\left(1-{}_nm_x^{(i,j)}\right)^{3\cdot{}_nP_x^{(i,j)}-{}_nD_x^{(i,j)}}$$

であるから、観測値 ${}_nD_x^{(i,j)}$ が与えられたときの ${}_nm_x^{(i,j)}$ の条件付き確率密度関数は、

$$\begin{aligned} p\left({}_nm_x^{(i,j)}\middle|{}_nD_x^{(i,j)}\right) &\propto p\left({}_nm_x^{(i,j)}\right)\cdot f\left({}_nD_x^{(i,j)}\middle|{}_nm_x^{(i,j)}\right) \\ &\propto \left({}_nm_x^{(i,j)}\right)^{{}_n\alpha_x^i+{}_nD_x^{(i,j)}-1}\left(1-{}_nm_x^{(i,j)}\right)^{{}_n\beta_x^i+3\cdot{}_nP_x^{(i,j)}-{}_nD_x^{(i,j)}-1} \end{aligned}$$

となる。

以上のことから、中央死亡率 ${}_nm_x^{(i,j)}$ の事後分布は再びベータ分布となり、

$${}_nm_x^{(i,j)}\left|\left({}_n\widetilde{D}_x^{(i,j)} = {}_nD_x^{(i,j)}\right)\right.\sim Beta\left({}_n\alpha_x^i + {}_nD_x^{(i,j)}, {}_n\beta_x^i + 3\cdot{}_nP_x^{(i,j)} - {}_nD_x^{(i,j)}\right)$$

であることが分かる。母数の推定値には、この事後分布の平均をあてることとし、

$$\begin{aligned} {}_nm_x^{(i,j)} &= \frac{{}_n\alpha_x^i + {}_nD_x^{(i,j)}}{{}_n\alpha_x^i + {}_n\beta_x^i + 3\cdot{}_nP_x^{(i,j)}} \\ &= \frac{{}_n\alpha_x^i + {}_n\beta_x^i}{{}_n\alpha_x^i + {}_n\beta_x^i + 3\cdot{}_nP_x^{(i,j)}}{}_nE_x^i + \frac{{}_nP_x^{(i,j)}}{{}_n\alpha_x^i + {}_n\beta_x^i + 3\cdot{}_nP_x^{(i,j)}}{}_n\widetilde{m}_x^{(i,j)} \end{aligned}$$

と推定した。

0 歳についても同様の推定をした。