3.1 Демографическое прогнозирование

«Прогнозы населения — научно обоснованная информация о будущих тенденциях изменения численности, параметров воспроизводства и структур населения на местном (региональном), национальном и глобальном уровнях» [25].

Демографические прогнозы можно классифицировать [25]

- по длине периода прогнозирования:

  краткосрочные прогнозы – до 5 лет;

  среднесрочные прогнозы – на 5–20 лет;

  долгосрочные прогнозы – на 20–50 лет;

- по количеству объектов прогнозирования:

  единичные (прогнозируется изменение одной переменной);

  множественные (прогнозируется изменение двух или более переменных);

- по типу представления прогнозируемой величины:

  точечные (прогнозируемая величина представлена одним числом);

  интервальные («прогнозируемая величина представляется в интервале показателей или в виде различных вариантов»);

- по методу построения:

  прогнозы, построенные математическим методом (прогнозируемая величина есть некоторая конкретная функция, зависящая от своего начального значения и времени);

  прогнозы, построенные на основе когортно-компонентного метода;

  прогнозы, построенные каузальным методом (прогнозируемая величина является зависимой переменной в регрессионном уравнении, связывающем ее с социально-экономическими показателями).

Этапами построения любого демографического прогноза являются следующие:

1. выбор модели, на основе которой будет строиться прогноз;
2. определение параметров выбранной модели, описывающих предстоящие изменения демографических показателей;
3. применение модели к «исходным демографическим показателям» (т. е. к реальным значениям демографических показателей, которые являются для модели начальными данными).

Математический (формульный) метод основан на использовании единой формулы, описывающей изменение общей численности населения (или какой-то отдельной группы населения). В рамках данного метода чаще всего рассматриваются модели экспоненциального и логистического роста.

Каузальный метод предназначен для прогнозирования отдельных демографических показателей. Модели, основанные на данном методе, являются эконометрическими и предназначены в том числе для уточнения моделей, основанных на методе компонент.

Единственным методом, позволяющим получить прогноз возрастно-половой структуры населения, является когортно-компонентный метод (метод передвижки по возрастам) [24].

3.2 Описание когортно-компонентного метода

Суть метода передвижки по возрастам состоит в следующем: численность мужчин и женщин в каждом возрасте x на начало года умножается на соответствующий коэффициент дожития, являющийся вероятностью дожить до возраста $x+1$ при условии уже состоявшегося дожития до возраста x. В результате получается численность населения возраста $x+1$ на начало уже следующего года. Такой расчет может быть применен к численностям всех возрастных групп на начало следующего года, кроме группы детей до 1 года.

Для расчета численности детей в возрасте до 1 года исчисляется среднегодовая численность женщин в каждом из репродуктивных возрастов и умножается на соответствующий возрастной коэффициент рождаемости, соответствующие произведения складываются по всем репродуктивным возрастам, и полученная общая численность распределяется на девочек и мальчиков с использованием коэффициента, характеризующего долю девочек среди всех новорожденных.

В стандартной схеме метода передвижки по возрастам к каждой из численностей возрастных групп также прибавляется сальдо миграции, соответствующее данному возрасту. В данном случае миграция исключена из рассмотрения, т. к. целью является изучение динамики собственного населения.

Наглядно и удобно метод передвижки по возрастам представляется с помощью матрицы Лесли. Пусть $n^s(t) \in R^{\omega+1}$ – вектор возрастного распределения женского ($s = f$) или мужского ($s = m$) населения, где ω – наибольший возможный возраст. Компонента n_x^s, $x = \overline{0,\omega}$, вектора n^s равна числу женщин (мужчин) возраста от x до $x+1$ лет. Тогда вектор возрастного распределения, например, женского закрытого (без учета миграции) населения удовлетворяет уравнению: $n^f\left(t+1\right) = L^f n^f(t)$, где L^f – матрица Лесли:

$$L^f = \begin{pmatrix} F_0^f & F_1^f & \ldots & F_{\omega-1}^f & F_\omega^f \\ P_0^f & 0 & \ldots & 0 & 0 \\ 0 & P_1^f & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & P_{w-1}^f & 0 \end{pmatrix}.$$

Здесь $P_x^f = \frac{L_{x+1}^f}{L_x^f}$ – коэффициент дожития (передвижки), $F_x^f = \frac{\delta}{2}\left(m_x + P_x^f m_{x+1}\right)\frac{L_1^f}{l_0}$, где L_x^f – численность женщин возраста x в стационарном населении таблицы смертности, m_x – возрастной коэффициент рождаемости, δ – доля девочек среди новорожденных, l_0 – так называемый корень таблицы смертности (в стандартном случае 100 тыс. человек). Формула $F_x^f = \frac{\delta}{2}\left(m_x + P_x^f m_{x+1}\right)\frac{L_1^f}{l_0}$ имеет место при предположениях о том, что все женщины, возраст которых на момент времени t составляет x лет, достигают возраста $x+1$ в момент времени $t+\frac{1}{2}$, кроме того, все смерти, произошедшие в году t, произошли в момент времени $t+\frac{1}{2}$. Очевидно, что часть элементов первой строки матрицы Лесли, соответствующих непродуктивным возрастам, равна нулю. С подробностями применения метода передвижки по возрастам можно ознакомиться в работе [7].

Модель вида $n^f\left(t+1\right) = L^f n^f(t)$ также называют моделью Лесли [26].

Аналогичная процедура может быть реализована и для мужского населения.

Если ввести в рассмотрение вектор-столбец n размерности $2\cdot(\omega+1)$, где в первых $\omega+1$ строках будет записано распределение женского населения по возрастам, а в следующих $\omega+1$ строках – распределение мужского населения, то $n\left(t+1\right) = Ln(t)$, где

$$L=\begin{pmatrix} F_0^f & \dots & \dots & F_\omega^f & 0 & \dots & \dots & 0 \\ P_0^f & 0 & \dots & 0 & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & P_{\omega-1}^f & 0 & 0 & \dots & \dots & 0 \\ F_0^m & \dots & \dots & F_\omega^m & 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & 0 & P_0^m & 0 & \dots & 0 \\ 0 & \dots & \dots & 0 & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & 0 & \dots & P_{\omega-1}^m & 0 \end{pmatrix} -$$

объединенная матрица Лесли. Здесь $P_x^m=\frac{L_{x+1}^m}{L_x^m}$, $F_x^m=\frac{1-\delta}{2}\left(m_x+P_x^f m_{x+1}\right)\frac{L_1^m}{l_0}$,

где L_x^m – численность мужчин возраста x в стационарном населении таблицы смертности.

Когортно-компонентый метод можно применять для прогнозирования численности населения, разбитого на m-летние возрастные группы при любом $m \geq 1$ (например, часто рассматриваются пятилетние возрастные группы). При этом, естественно, шаг прогнозирования в 1 год нужно заменить на шаг, составляющий m лет.

Отметим, что выше описан когортно-компонентный метод прогнозирования с постоянной матрицей Лесли, что соответствует гипотезе о сохранении фиксированных демографических характеристик в течение всего периода прогнозирования. На практике чаще всего прогнозирование осуществляется на основе каких-либо гипотез относительно будущих показателей рождаемости и смертности, что соответствует использованию различных матриц Лесли на каждом шаге (прогнозирование с переменной матрицей Лесли). Так, например, Росстат чаще всего дает 3 варианта прогноза: низкий, средний и высокий. Низкий вариант основан на «экстраполяции существующих демографических тенденций»; высокий вариант «ориентирован на достижение целей, определенных в Концепции демографической политики Российской Федерации на период до 2025 года» и потому является наименее реалистичным; средний вариант является комбинацией низкого и высокого вариантов, а именно, учитывает как сложившиеся ранее демографические тенденции, так и «принимаемые меры демографической политики» [90]. Распространенным является также построение многовариантных прогнозов (так называемый сценарный подход), см. например [27].

3.3 Когортно-компонентный метод. Математический аспект

3.3.1 Основные свойства оператора Лесли

Решение уравнения $n^f(t+1)=L^f n^f(t)$, соответствующее начальному распределению по возрастам $n^f(0)$, имеет вид $n^f(t)=\left(L^f\right)^t n^f(0)$, где $\left(L^f\right)^t$ — степень постоянной матрицы L^f. Будем далее опускать верхний индекс f, не забывая о том, что мы рассматриваем динамику численности одного пола. Переход к численности населения обоих полов очевиден.

Матрица L определяет линейный оператор в n-мерном евклидовом пространстве, называемый оператором Лесли. Очевидно, матрица L является неотрицательной, поэтому оператор Лесли действует из положительного ортанта в положительный ортант.

Асимптотическое поведение решений уравнения зависит от спектральных свойств оператора Лесли. Изучение этих свойств основано на применении теоремы Перрона-Фробениуса, которая применима к неразложимым матрицам. Для неразложимости матрицы Лесли необходимо и достаточно, чтобы $F_\omega^f \neq 0$ [26]. Это условие означает, что в качестве ω выступает не наибольший возможный возраст, а наибольший репродуктивный возраст, поэтому далее будем рассматривать динамику численностей младших и репродуктивных групп населения. Здесь важно отметить, что в модели отсутствует эффект лимитирования по общей численности населения, поэтому динамика численности пострепродуктивных групп не оказывает никакого влияния на динамику численности младших и репродуктивных групп населения. Численности пострепродуктивных групп элементарно определяются в любой момент времени через численность последней репродуктивной группы и потому могут быть рассмотрены отдельно.

Нетривиальное предельное распределение $n = \lim_{t \to \infty} L^t n(0)$ существует тогда, когда максимальное вещественное собственное число матрицы L (его наличие гарантируется теоремой Перрона-Фробениуса) равно 1. Если же максимальное собственное число матрицы L меньше 1, то население с течением времени вырождается; если больше 1 – неограниченно растет.

Перейдем к вопросу о существовании устойчивых траекторий в динамической системе $n(t+1) = Ln(t)$. Равновесие, т. е. решение уравнения $n^* = Ln^*$, является собственным вектором матрицы L, соответствующим максимальному собственному числу, равному единице. Показано, что в случае если равновесие существует, то оно локально устойчиво; асимптотическая устойчивость в этом случае отсутствует. Также для модели Лесли доказана невозможность существования хаотических режимов [26].

3.3.2 Модель Лесли как дискретизация непрерывной модели динамики структуры популяции

Опишем непрерывную модель динамики структуры популяции [23]. В данном случае вводится возрастная плотность численности, т. е. функция $x(\tau,t)$ такая, что интеграл $\int_{\tau_1}^{\tau_2} x(\tau,t)\,d\tau$ представляет собой численность населения в возрасте от τ_1 до τ_2 в момент времени t. Кроме того, определены коэффициенты рождаемости и смертности $b(\tau,t)$ и $d(\tau,t)$ соответственно, такие, что в момент времени t численность новорожденных, рожденных населением в возрасте от τ_1 до τ_2, и численность населения, умершего в возрасте от τ_1 до τ_2, вычисляются как $\int\limits_{\tau_1}^{\tau_2} b(\tau,t)\,x(\tau,t)\,d\tau$ и $\int\limits_{\tau_1}^{\tau_2} d(\tau,t)\,x(\tau,t)\,d\tau$ соответственно.

Система уравнений, описывающая динамику структуры популяции, состоит из так называемых уравнения рождаемости и уравнения выживаемости. Уравнение рождаемости, имеющее вид $x(0,t)=\int\limits_{0}^{\infty} b(\tau,t)\,x(\tau,t)\,d\tau$, не требует дополнительных пояснений.

Перейдем к выводу уравнения выживаемости. Зафиксируем некоторые τ и t. Численность населения в возрасте от τ до $\tau+\Delta\tau$ в момент времени $t+\Delta t$ равна численности населения в возрасте от $\tau-\Delta t$ до $\tau+\Delta\tau-\Delta t$ в момент времени t за вычетом населения, умершего за период времени Δt. При малых приращениях $\Delta\tau$ и Δt данное утверждение можно с точностью до линейных членов можно выразить следующим уравнением:

$$x(\tau,t+\Delta t)\Delta\tau = x(\tau-\Delta t,t)\Delta\tau - d(\tau,t)\,x(\tau,t)\Delta\tau\Delta t.$$

Разделив данное равенство на $\Delta\tau$ и Δt и вычтя из обеих частей $x(\tau,t)$, получим $\dfrac{x(\tau,t+\Delta t)-x(\tau,t)}{\Delta t}+\dfrac{x(\tau,t)-x(\tau-\Delta t,t)}{\Delta t}=-d(\tau,t)\,x(\tau,t)$. Переходя в последнем равенстве к пределу при $\Delta t\to 0$, получим уравнение выживаемости, имеющее вид $\dfrac{\partial x}{\partial t}+\dfrac{\partial x}{\partial \tau}=-d(\tau,t)\,x(\tau,t)$.

Итак, имеем систему

$$\begin{cases} \dfrac{\partial x}{\partial t}+\dfrac{\partial x}{\partial \tau}=-d(\tau,t)\,x(\tau,t), \\ x(0,t)=\int\limits_{0}^{\infty} b(\tau,t)\,x(\tau,t)\,d\tau \end{cases}$$

при начальном условии $x(\tau,0)=\varphi(\tau)$.

Стандартным является условие «согласования» $x(0,0)=\varphi(0)=\int\limits_0^\infty b(\tau,0)\varphi(\tau)d\tau$, тем не менее, возможно более естественным представляется рассматривать решение $x(\tau,t)$, разрывное при $t=0$, считая, что

$$x(0,t)=\begin{cases}\varphi(0),\ t=0,\\ \int\limits_0^\infty b(\tau,t)x(\tau,t)d\tau,\ t>0.\end{cases}$$

Отметим, что рассматриваемая задача не является классической задачей математической физики, поскольку в граничное условие $x(0,t)=\int\limits_0^\infty b(\tau,t)x(\tau,t)d\tau$ входит все искомое распределение $x(\tau,t)$ [23].

Покажем теперь, что дискретизация данной непрерывной модели приводит к модели Лесли, рассмотренной выше. Проведем дискретизацию по переменным τ и t с одинаковым шагом, равным единице.

Начнем с дискретизации по τ: $\tau_0=0$, $\tau_1=1$, $\tau_2=2,\ldots,\ \tau_n=\omega$. Введем следующие обозначения $x_0(t)=x(\tau_0,t)$, $x_1(t)=x(\tau_1,t)$ и т.д. Замена интеграла на конечную сумму в уравнении рождаемости приводит нас к следующему равенству:

$$x(0,t)=\sum_{i=1}^{n}b(\tau_i,t)x_i(t)=\sum_{i=1}^{n}b_i x_i(t). \tag{14}$$

Далее, заменим производную $\frac{\partial x}{\partial \tau}$ в уравнении выживаемости простейшим конечно-разностным отношением, получим

$$\frac{dx_1}{dt}+x_1(t)-x_0(t)=-d(0,t)x_0(t)=-d_0(t)x_0(t), \tag{15}$$

$$\ldots,$$

$$\frac{dx_n}{dt} + x_n(t) - x_{n-1}(t) = -d\left(n-1,t\right)x_{n-1}(t) = -d_{n-1}(t)x_{n-1}(t).$$

Поскольку численность группы в возрасте от τ_i до τ_{i+1} выражается как $\int_{\tau_i}^{\tau_{i+1}} x\left(\tau,t\right)d\tau$, что можно приближенно вычислить как $x\left(\tau_i,t\right)\left(\tau_{i+1}-\tau_i\right) = x_i(t)$, $x_i(t)$ есть численности соответствующих возрастных групп (в возрасте от i до $i+1$).

Подставляя уравнение (14) в первое из уравнений (15) и переписывая прочие уравнения (15), получаем

$$\frac{dx_1}{dt} = -x_1(t) - d_0(t)x_0(t) + \sum_{i=1}^{n} b_i x_i(t),$$

$$\frac{dx_2}{dt} = -x_2(t) + x_1(t) - d_1(t)x_1(t),$$

$$\ldots,$$

$$\frac{dx_n}{dt} = -x_n(t) + x_{n-1}(t) - d_{n-1}(t)x_{n-1}(t).$$

Далее, проводя дискретизацию по t, имеем

$$x_1(k+1) = -d_0(k)x_0(k) + \sum_{i=1}^{n} b_i(k)x_i(k),$$

$$x_2(k+1) = x_1(k) - d_1(k)x_1(k) = \left(1-d_1(k)\right)x_1(k),$$

$$\ldots,$$

$$x_n(k+1) = x_{n-1}(k) - d_{n-1}(k)x_{n-1}(k) = \left(1-d_{n-1}(k)\right)x_{n-1}(k).$$

Итак, с точностью до обозначений получена модель Лесли, точнее $n\left(t+1\right) = L(t)n(t)$, где $t = 0,\ 1,\ 2,\ \ldots$,

$$n(t) = \begin{pmatrix} x(0,t) \\ x(1,t) \\ \ldots \\ x(n,t) \end{pmatrix},\ L = \begin{pmatrix} b(0,t) & b(1,t) & \ldots & b(n-1,t) & b(n,t) \\ 1-d(0,t) & 0 & \ldots & 0 & 0 \\ 0 & 1-d(1,t) & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 1-d(n-1,t) & 0 \end{pmatrix}.$$

Данное уравнение решается при начальном условии

$$n(0)=n_0=\begin{pmatrix}\varphi(0)\\ \varphi(1)\\ \ldots\\ \varphi(n)\end{pmatrix}.$$

Таким образом, полученная дискретная форма модели динамики структуры популяции при зависящих от времени коэффициентах рождаемости и смертности $b(\tau,t)$ и $d(\tau,t)$ приводит к когортно-компонентному методу прогнозирования с переменной матрицей Лесли.

Отметим, что в случае не зависящих от времени коэффициентов b и d решение системы

$$\begin{cases}\dfrac{\partial x}{\partial t}+\dfrac{\partial x}{\partial \tau}=-d(\tau)x(\tau,t),\\ x(0,t)=\displaystyle\int_0^{\infty} b(\tau)x(\tau,t)d\tau\end{cases}$$

при начальном условии $x(\tau,0)=\varphi(\tau)$ может быть найдено аналитически.

Замена переменных $u=\tau-t$, $v=t$ $(\tau=u+v,\ t=v)$ сводит уравнение выживаемости к обыкновенному дифференциальному уравнения первого порядка $\dfrac{dx}{dv}=-d(u+v)\cdot x(u+v,v)$, решением которого является функция $x(u,v)=\Omega(u)\cdot e^{-\int_0^{u+v} d(\xi)d\xi}$, где Ω – произвольная функция. Итак, общее решение уравнения выживаемости имеет вид $x(\tau,t)=\Omega(t-\tau)e^{-\int_0^{\tau} d(\xi)d\xi}$.

Подстановка $x(\tau,t)=\Omega(t-\tau)e^{-\int_0^{\tau} d(\xi)d\xi}$ в уравнение рождаемости $x(0,t)=\int_0^{\infty} b(\tau,t)x(\tau,t)d\tau$ и начальное условие $x(\tau,0)=\varphi(\tau)$ позволяет однозначно определить функцию Ω, а именно, при отрицательных значениях аргумента Ω выражается следующим образом: $\Omega(-\tau)=\tilde{\varphi}(\tau)=\varphi(\tau)e^{\int_0^\tau d(\xi)d\xi}$, $\tau \geq 0$, при положительных значениях аргумента Ω является решением интегрального уравнения $\Omega(t)=\int_0^t K(\tau)\Omega(t-\tau)d\tau+\psi(t)$, $t \geq 0$, где $K(\tau)=b(\tau)e^{-\int_0^\tau d(\xi)d\xi}$, $\psi(t)=\int_t^\infty K(\tau)\Omega(t-\tau)d\tau=\int_t^\infty K(\tau)\tilde{\varphi}(t-\tau)d\tau$.

Интегральное уравнение решается при помощи преобразования Лапласа. По теореме о свертке $\Omega^*=K^*\Omega^*-\psi^*$, где через функции со знаком «*» обозначены изображения по Лапласу соответствующих функций. Таким образом, функция Ω находится из уравнения $\Omega^*=\dfrac{\psi^*}{1-K^*}$. В [23] показано, что $\Omega(t)=\sum_{i=1}^{\infty} c_i e^{\lambda_i t}$, где λ_i – полюсы функции $\Omega^*(s)$ комплексной переменной, а $c_i=\underset{\lambda_i}{res}\,\Omega^*(s)$ – ее вычеты в соответствующих полюсах. Окончательно решение задачи может быть представлено в виде $x(\tau,t)=e^{-\int_0^\tau d(\xi)d\xi}\cdot\sum_{i=1}^{\infty} c_i e^{\lambda_i(t-\tau)}$.

Таким образом, решение системы, описывающей непрерывный аналог метода «передвижки по возрастам» с постоянной матрицей Лесли, может быть найдено аналитически. Однако идентификация коэффициентов рождаемости и смертности $b(\tau)$ и $d(\tau)$ по статистическим данным снова неминуемо ведет к дискретизации, что свидетельствует о том, что на практике оправдано использование модели Лесли. Тем не менее, некоторыми авторами, см. например [10], используется непрерывная модель структуры популяции с определением функций $b(\tau)$ и $d(\tau)$ через интерполяцию статистических данных.