Е. М. АНДРЕЕВ

О СВЯЗИ РЕАЛЬНОГО И ГИПОТЕТИЧЕСКОГО ПОКОЛЕНИЙ

Данная статья посвящена изучению взаимосвязи реальных демографических процессов и их отражения системой демографических показателей. Изучение этой взаимосвязи мы проведем на примере процесса смертности, однако сделанные выводы могут быть распространены и на иные демографические процессы.

Основное внимание уделено проблеме связи реального и гипотетического поколений, а также анализу возможных точек зрения на определение гипотетического поколения. Целью статьи является количественное измерение отличий показателей таблиц для реальных и гипотетических поколений, которые весьма существенны в условиях быстро меняющейся демографической ситуации. Следует отметить, что связи и различия, количественному анализу которых посвящена статья, на качественном уровне неоднократно рассматривались в демографической литературе.

1. Исходная математическая модель населения

Мы будем исходить из непрерывной модели популяции. Рассмотрим вначале функцию $l(x, t)$, равную вероятности дожития до точного возраста x для лица, родившегося в момент времени t. Функция $l(x, t)$ определена при $x \geqslant 0$ и при всех t, т. е. на верхней половине координатной плоскости (t, x), именуемой демографической сеткой или плоскостью демографических событий. Рассмотрим также функцию $\rho(t)$ — плотность рождений. Функция $\rho(t)$ определяется следующим образом. Обозначим через $N(t_1, t_2)$ число рождений на отрезке времени $[t_1, t_2]$. Плотностью рождений называется такая функция $\rho(t)$, что для любых $t_1 \leqslant t_2$

$$N(t_1 t_2) = \int_{t_1}^{t_2} \rho(t)\,dt \quad . \tag{1}$$

Основным допущением непрерывной модели является предположение, что функция $\rho(t)$ непрерывна, а функция $l(x, t)$ непрерывно дифференцируема, т. е. имеет непрерывные производные по x и по t.

Оба эти допущения, естественно, могут не выполняться в реальной популяции; более того, функция $\rho(t)$ заведомо не непрерывна. Это, естественно, снижает ценность непрерывной модели и ставит под сомнение адекватность выводов, сделанных на базе этой модели, реальной ситуации. Обширная практика демографических расчетов доказывает беспочвенность таких сомнений, однако в интересах формальной строгости и логической завершенности полезно специально рассмотреть вопрос о степени близости непрерывной модели к реальной ситуации.

В популяции, состоящей из одного человека, родившегося в момент времени t_0, функция $\rho(t)$ должна обладать двумя свойствами: если отрезок $[t_1, t_2]$ не содержит точки t_0, то $\int_{t_1}^{t_2} \rho(t)dt = 0$, но если точка t_0 лежит на отрезке $[t_1, t_2]$, т. е. $t_1 \leqslant t_0 \leqslant t_2$, то $\int_{t_1}^{t_2} \rho(t)dt = 1$. Эти соотношения есть не что иное, как определение дельта-функции $\delta(t-t_0) = \rho(t)$ ции[1] $\delta(t-t_0)$, т. е. $\rho(t) = \delta(t-t_0)$.

Если же в популяции N членов, а i-й ее член родился в момент t_i, то

$$\rho(t) = \sum_{i=1}^{N} \delta(t - t_i) \quad . \tag{2}$$

Функцию $\delta(t)$ именуют обобщенной функцией. Она может быть с любой точностью приближена непрерыв-

[1] Основные сведения о дельта-функции см. в работах: *Я. Микусинский* и *Р. Сикорский.* Элементарная теория обобщенных функций. Вып. 1. М., Изд-во иностранной литературы, 1959; вып. 2. М., Изд-во иностранной литературы, 1963; *Н. Я. Виленкин* и *др.* Функциональный анализ. Под ред. С. Г. Крейна. М., «Наука», 1964.

ными функциями в следующем смысле: для любого (сколь угодно малого) $h>0$ существует функция $\varphi_h(t)$ такая, что $\varphi_h(t)=0$ при $|t|>h$, и $\int\limits_{-h}^{+h}\varphi_h(t)dt=1$.

Ясно, что $\int\limits_{t_1}^{t_2}\varphi_h(t)dt$ равен 0, если отрезок $[-h, +h]$ не имеет общих точек с отрезком $[t_1, t_2]$, и равен 1, если отрезок $[-h, +h]$ целиком лежит в отрезке $[t_1, t_2]$. Заменим в формуле (2) $\delta(t)$ на $\varphi_h(t)$:

$$\rho_h(t)=\sum_{i=1}^{N}\varphi_h(t-t_i)\ . \qquad (3)$$

Если выбрать значение h столь малым, чтобы оно было, во-первых, меньше той точности, с какой мы измеряем время, а, во-вторых, чтобы среднее число рождений на отрезке длины $6h$ было незначительно по сравнению с числом рождений на рассматриваемых отрезках $[t_1, t_2]$, то можно добиться, чтобы для любого отрезка $[t_1, t_2]$, длина которого больше некоторой фиксированной величины D, и для любого $\varepsilon>0$ выполнялось условие

$$\left|\int\limits_{t_1}^{t_2}(\rho(t)-\rho_h(t))dt\right|<\varepsilon\ . \qquad (4)$$

Величины D и ε могут быть выбраны сколь угодно малыми, т. е. с точки зрения вычисления $N(t_1, t_2)$ и вообще любого интеграла $\int\limits_{t_1}^{t_2}f(t)\rho(t)dt$, где функция $|f(t)|<1$, функция $\rho(t)$ может быть приближена $\rho_h(t)$ с любой точностью, а значит, задавшись фиксированной точностью, мы можем заменить не непрерывную функцию $\rho(t)$ непрерывной (или даже дифференцируемой) функцией $\rho_h(t)$. В дальнейшем мы будем предполагать, что исходная функция $\rho(t)$ непрерывна.

Несколько сложнее обстоит дело с функцией $l\ (x,\ t)$. Отметим прежде всего, что численность живущих в реальной популяции меняется не непрерывно и что число смертей на отрезке возрастов $[x_1, x_2]$ может существенно отличаться от своего математического ожидания, а, следовательно, отношение числа событий к численности населения, называемое частотой, может быть отлично от вероятности события. Однако в силу действия закона больших чисел это отличие тем меньше, чем многочисленнее популяция[2], и в достаточно многочисленных популяциях этим различием можно пренебречь. В силу этого мы можем допустить, что в рассматриваемой модели населения вероятность события совпадает с его частотой. Тогда число доживших до возраста x из совокупности родившихся за период $[t_1, t_2]$ составит

$$\int_{t_1}^{t_2} l(x, t)\rho(t)\,dt = S(x; t_1 t_2) \quad (5)$$

Подобрать непрерывно дифференцируемую по x и t функцию $l\ (x,\ t)$ так, что в реальной популяции

$$\left|S(x; t_1, t_2) - \int_{t_1}^{t_2} l(x, t)\rho(t)\,dt\right| < \varepsilon ,$$

можно с помощью тех же допущений, что и функцию $\rho\ (t)$. При этом величина $l\ (x,\ t)$ будет характеризовать перспективы дожития лиц, родившихся не в момент t, а на некотором интервале $(t-h, t+h)$; величина $s(x;\ t_1, t_2)$, вычисленная по формуле (5), будет характеризовать не точную численность в возрасте x, а некоторую среднюю численность в интервале возрастов $(x-h,\ x+h)$.

В реальной статистике вероятность дожития вычисляется не для лиц, родившихся в данный момент t или в данный малый интервал времени, но для рожденных в течение, достаточно протяженного временного отрезка $[t_1,\ t_2]$. Такую совокупность называют реальным поколением родившихся в период t_1, t_2. В нашей модели эта вероятность, обозначим ее $l(x;\ t_1,\ t_2)$, находится по формуле

$$l(x; t_1, t_2) = \frac{\int_{t_1}^{t} \rho(t)\, l(x, t)\,dt}{\int_{t_1}^{t} \rho(t)\,dt}, \quad (6)$$

[2] Точную формулировку закона больших чисел см., например, в учебнике для втузов Е. С. Вентцель «Теория вероятностей» (М., Физматгиз, 1962).

или

$$l(x; t_1, t_2) = \frac{S(x;\ t_1, t_2)}{N(t_1, t_2)}. \tag{6а}$$

Исходя из функции $l(x, t)$ или $l\ (x;\ t_1,\ t_2)$, можно развернуть целую систему интервальных показателей, т. е. функций, характеризующих смертность на тех или иных возрастных интервалах. Выразим вероятность для лица, дожившего до точного возраста x, дожить до точного возраста $x+v$

$$_vp(x) = \frac{l(x+v)}{l(x)}. \tag{7}$$

Для реального поколения эта формула примет вид

$$_vp(x) = \frac{S(x+v;\ t_1, t_2)}{S(x;\ t_1, t_2)}. \tag{7а}$$

Вероятность для лица, дожившего до точного возраста x, умереть, не дожив до возраста $x+v$, составит

$$_vq(x) = 1 - {_vp(x)}. \tag{8}$$

Обозначим через $M(x,\ x+v;\ t_1,\ t_2)$ число смертей в интервале возрастов $(x,\ x+v)$ у лиц, родившихся в интервале $(t_1,\ t_2)$. Тогда

$$_vq(x) = \frac{M(x, x+v;\ t_1, t_2)}{S(x;\ t_1, t_2)}; \tag{8а}$$

$$_vd(x) = l(x) - l(x+v) - \tag{9}$$

вероятность для новорожденного умереть в интервале возрастов $(x,\ x+v)$;

$$_vL(x) = \int\limits_x^{x+v} l(y)\,dy - \tag{10}$$

среднее число лиц, доживших до возраста y, из интервала $(x,\ x+v)$, помноженное на длину интервала.

Интервальные показатели не только характеризуют уровень смертности данного возраста, но и позволяют с известной точностью восстанавливать функцию $l\ (x)$. Дело в том, что во многих практических ситуациях прямое вычисление $l\ (x)$ затруднено или невозможно. Особенно это касается открытых популяций, т. е. популяций с миграцией, выход из которых происходит не только по причине смерти, а вход в которые возможен не только вследствие рождения. При изучении иных демографических процессов (брачности, плодовитости) повсеместно приходится сталкиваться с открытыми популяциями. В целом аналогичная ситуация встречается и при изучении смертности по причинам.

В условиях миграции число живущих в возрасте x из числа родившихся в интервале (t_1, t_2), т. е. $S\,(x; t_1, t_2)$, не является подсовокупностью родившихся. Возможно, $S(x;\ t_1,\ t_2) > N(t_1,\ t_2)$ или вообще $S(x+v;\ t_1,\ t_2) > S(x; t_1, t_2)$, и, следовательно, даже интервальные показатели не могут быть вычислены непосредственно.

В открытой популяции, строго говоря, чистый, прямой, смысл имеет показатель $\mu\,(x,\ t)$ — сила смертности. Величина $\mu\ (x)$ (мы упускаем время рождения для простоты записи) определяется следующим образом. Предполагается, что вероятность смерти в возрастном интервале $(x,\ x+\Delta x)$ для лиц, доживших до возраста x, пропорциональна Δx с точностью до бесконечно малых высшего порядка, т. е.

$$_{\Delta x}q\,(x) = c\Delta x + o(\Delta x), \tag{11}$$

где $o\,(\Delta x)$ стремится к 0 при $\Delta x \longrightarrow 0$ быстрее, чем Δx: $\lim\limits_{\Delta x\to 0} \frac{o\,(\Delta x)}{\Delta x} = 0$. Коэффициент пропорциональности в возрасте x и есть $\mu\ (x)$.

Отсюда

$$l(x) - l(x+\Delta x) = l(x)\,\mu(x)\,\Delta x + l(x)\,o(\Delta x)$$

или

$$-\frac{l(x+\Delta x) - l(x)}{\Delta x} = l(x)\,\mu(x) + \frac{l(x)\,o(\Delta x)}{\Delta x}.$$

Переходя к пределу при $\Delta x \longrightarrow 0$, имеем

$$-l'(x) = l(x)\,\mu(x). \tag{12}$$

Решая это уравнение, находим, что

$$\frac{d}{dx}\ln l(x) = -\mu(x), \tag{13}$$

или же

$$l(x) = e^{-\int\limits_0^x \mu(y)dy}, \tag{14}$$

так как $l(0) = 1$.

Преимущество силы смертности $\mu(x)$ перед другими показателями состоит в том, что она описывает процесс на столь малом интервале, что в условиях сложного процесса (например, смертность в открытой популяции) на этом интервале *не может* произойти более одного события, т. е. вероятность того, что человек и мигрирует и умрет за интервал $(x, x+\Delta x)$, есть бесконечно малая порядка выше Δx.

В условиях открытой популяции все иные показатели, кроме μ, имеют смысл условной характеристики уровня смертности. Они описывают жизнь модельного замкнутого (постоянного) населения, которое существует на протяжении всей своей жизни в тех условиях, при том уровне смертности, который сложился на данной территории. Чем больше интервал, тем больше условность показателя, чем меньше — тем ближе модель к реальности, и лишь $\mu\ (x)$ — абсолютно адекватный показатель.

Однако $\mu\ (x)$ не может быть вычислен непосредственно. Как следует из определения,

$$\mu(x) = \lim_{v \to 0} \frac{{}_v q(x)}{v}, \tag{15}$$

т. е. равен пределу отношения ${}_v q(x)/v$ при v стремящемся к 0. Практически переход к нулевому пределу можно заменить вычислением при достаточно малом v (которое может оказаться вполне зримой величиной).

Другой подход к $\mu(x)$ возможен со стороны табличного коэффициента смертности:

$$ {}_v m(x) = \frac{{}_v d(x)}{{}_v L(x)}\ . \tag{16}$$

По определению ${}_v d\ (x)$ и ${}_v L\ (x)$ имеем

$$ {}_v m(x) = \frac{l(x) - l(x+v)}{\int\limits_x^{x+v} l(y)\,dy} = \frac{l(x) - l(x+v)}{l(x+\xi(v)\cdot v)\cdot v}$$

где $0 \leqslant \xi(v) \leqslant 1$; равенство $\int\limits_x^{x+v} l(y)\,dy = l(x+\xi(v)\ v)\cdot v$ следует из теоремы о среднем. Переходя к пределу при $v \longrightarrow 0$, увидим, что

$$\mu(x) = \lim_{v \to 0} {}_v m(x). \tag{17}$$

Однако $\mu(x)$ и ${}_vm(x)$ не следует смешивать полностью. Отметим, что ${}_vd(x) \equiv l(x) - l(x+v) = -\frac{\partial}{\partial x}\int\limits_x^{x+v} l(y)\quad dy =$

$= -\frac{\partial}{\partial x}\,{}_vL(x)$. Промежуточное равенство следует из теоремы о том, что производная интеграла по верхнему пределу равна значению подинтегральной функции в этой точке, а по нижнему — ее значению со знаком минус.

Итак,

$$\frac{\frac{\partial}{\partial x^v} L(x)}{{}_vL(x)} = \frac{\partial}{\partial x}\ln{}_vL(x) = -\frac{{}_vd(x)}{{}_vL(x)} = -{}_vm(x), \tag{18}$$

а это означает, что

$$ {}_vL(x) = {}_vL(0)\,e^{-\int\limits_0^x {}_vm(y)dy}. \tag{19}$$

Полученные формулы (18) и (19) можно интерпретировать следующим образом. Пусть для простоты $v=1$. Рассмотрим поколение, родившееся за период [—1,0]. Допустим, что на отрезке [—1,0] плотность рождений постоянна, а общее число родившихся равно N. Ясно, что $\rho(t) = N; -1 \leqslant t \leqslant 0$. Предположим[3], что функция $l(x, t)$ постоянна по t на отрезке [—1,0]. В этих условиях число лиц, доживших до точного возраста x, равно $Nl(x)$. Вычислим, чему равно число живущих в момент z:

$$\hat{S}(z) = \int\limits_z^{z+1} l(x)\rho(z-x)\,dx = N_1L(z). \tag{20}$$

Таким образом, ${}_1L(x)$ в условиях постоянства плотности рождения на отрезке $[t, t+1]$ равно доле лиц из начальной совокупности родившихся, доживших до момента времени $z = x - t + 1$.

Вернемся к нашему примеру. Из равенства (18) следует, что

$$ {}_1L(x) - {}_1L(x+\Delta x) = {}_1m(x)\Delta x L(x) + o(\Delta x).$$

---

[3] Эти допущения являются стандартными гипотезами для реального поколения, принятыми при построении таблиц смертности.

Заменим x формально на z и, используя формулу (20), получим

$$\hat{S}(z)-\hat{S}(z+\Delta z)={}_1m(z)\,\Delta z\hat{S}(z)+o(\Delta z).$$

Таким образом, ${}_1m(z)\Delta z$ есть с точностью до бесконечно малой высшего порядка вероятность смерти на отрезке времени $[z, z+\Delta z]$. Итак, в отличие от пары $\mu(x)$, $l(x)$, характеризующей процесс вымирания по возрасту, ${}_vm(x)$, ${}_vL(x)$ характеризуют процесс вымирания как бы по времени.

В рамках того же примера отметим, что

$$ {}_1m(x)=\frac{M(x, x+1; -1, 0)}{\hat{S}(z; -1, 0)} \qquad (21)$$

при $z=x$.

Как мы уже говорили выше, миграция затрудняет вычисление и интервальных показателей, однако с помощью разумных гипотез о силе смертности и плотности миграции можно с известной достоверностью и точностью исчислять и интервальные показатели ${}_vq(x)$, ${}_vm(x)$ и другие, описывающие смертность не на бесконечно малом, а на конечном интервале v[4].

2. Гипотетическое поколение. Связь смертности реальных и гипотетических поколений

Всюду до сих пор мы рассматривали реальные поколения. Наряду с ними в демографии повсеместно изучается так называемое гипотетическое поколение. Более того, именно для гипотетического поколения строилось и строится большинство таблиц смертности.

Представим себе поколение людей, родившихся и проживших всю жизнь по законам, характерным для сегодняшней демографической ситуации. Такая, быть может, реально не существующая, совокупность людей и называется гипотетическим поколением сегодняшнего

---

[4] См., например: *В. В. Паевский.* Об измерении смертности мигрирующих масс населения. — «Труды Демографического института АН СССР», т. 1. Л., Изд-во АН СССР, 1934; *Ю. А. Корчак-Чепурковский.* О методах изучения воспроизводства населения. — В его кн. «Избранные демографические исследования». М., «Статистика», 1970, с. 73—86.

дня. Если ограничиться смертностью, то гипотетическим поколением следует назвать воображаемую совокупность людей, вся жизнь которых прошла в условиях смертности данного момента времени.

В тот период истории, когда возникла идея гипотетического поколения, смертность была почти неизменна с точностью до отдельных резких колебаний смертности, связанных с голодом и эпидемиями; так что смертность гипотетического поколения довольно точно отражала смертность реальных поколений и в случае выбора подходящего момента времени позволяла устранить влияние преходящих явлений. С развитием человечества в одних странах раньше, в других позже начался активный процесс снижения смертности, уровень смертности ныне ни в одной стране не постоянен. Однако модель гипотетического поколения вовсе не отпала, изменилась лишь интерпретация показателей, рассчитанных на ее основе.

Мы не будем приводить здесь аргументы в защиту гипотетического поколения. Укажем лишь, что иной способ характеристики смертности данного момента времени не известен и что при изучении динамики смертности за короткий промежуток времени, близкий к моменту наблюдения, так или иначе приходится пользоваться конструкцией, родственной гипотетическому поколению.

Наконец, при анализе смертности открытой популяции реальное поколение по году рождения само весьма условно, ибо не является постоянной совокупностью. Смертность в таком поколении характеризует условное население, всю жизнь прожившее в данных условиях.

Ниже мы дадим строгое определение гипотетического поколения.

Рассмотрим какой-нибудь треугольник ABC на демографической сетке (см. рис. 1).

Любые показатели, характеризующие демографические процессы как функцию возраста, могут быть расположены вдоль вертикальной линии жизни AC и относиться к реальному поколению людей, родившихся в некоторый момент t. Если мы возьмем конкретное поколение родившихся в какой-то определенный момент t_0 или в течение определенного отрезка времени от t_1 до t_2, то функция, которая в общем виде записывается как $f(x, t)$, для данного поколения приобретает вид $f(x, t_0)$, или $f(x; t_1, t_2)$. Выражаясь математически, функция

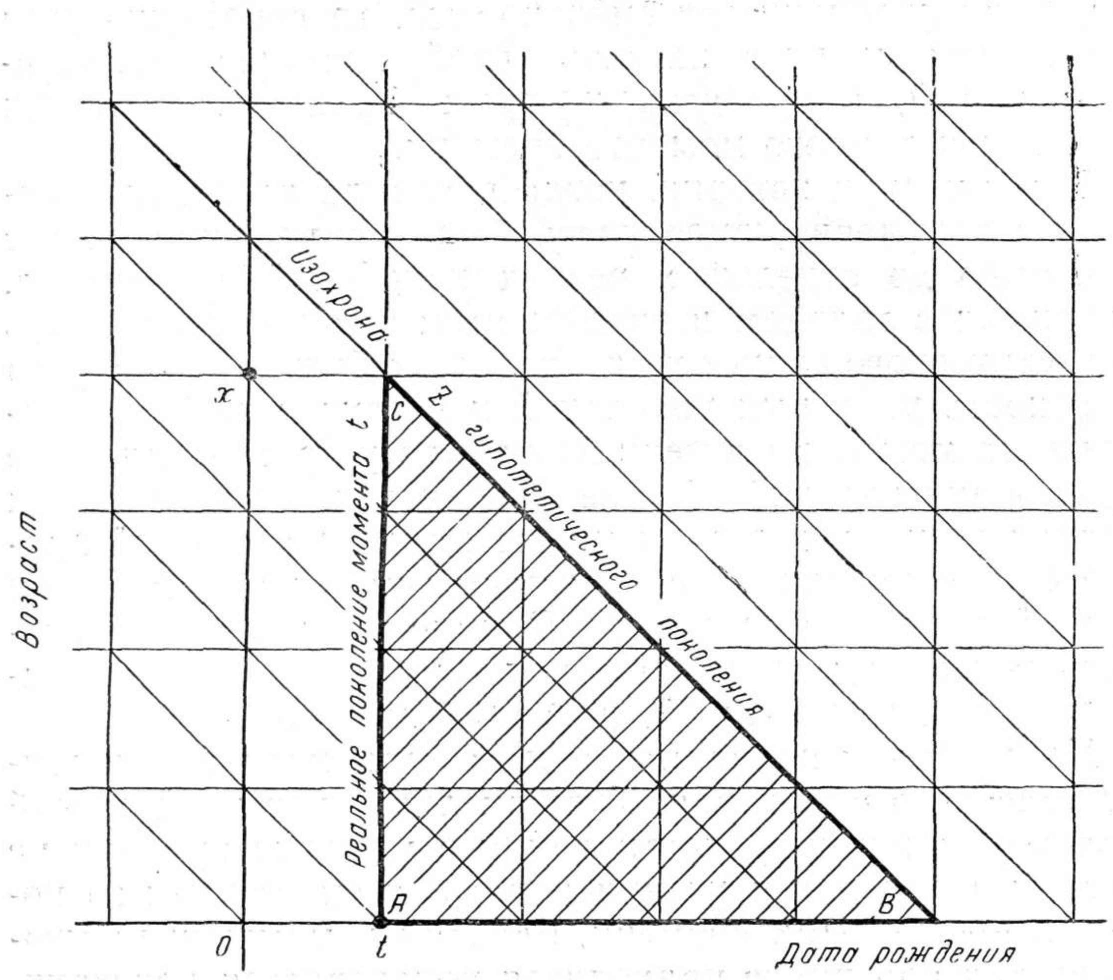

Рис. 1.

$f(x, t)$ ограничена на прямую $t = t_0$ или на полосу $t_1 \leqslant t \leqslant t_2$.

Не все показатели, о которых здесь идет речь, могут быть расположены также и вдоль линии наблюдения, или изохроны, — гипотенузы треугольника BC. Эти показатели также характеризуют интенсивность демографических процессов в возрасте x, но уже не для конкретного поколения лиц, родившихся в фиксированный момент t_0, а для тех, кто в некоторый момент z находится в возрасте x, т. е. для тех, кто родился в момент $t_0 = z_0 - x$, свой для каждого x. Они описываются функцией $f(x, z_0 - x)$, которую мы обозначим просто как $f(x; z_0)$. Тем самым мы ограничили все ту же функцию на изохрону $z_0 = x + t_0$. Ее можно ограничить также на полосу изохрон $z_1 \leqslant z \leqslant z_2$ (если речь идет не о моменте, а о периоде наблюдения): $f(x; z_1, z_2)$.

Ограничение функции на полосу означает, что рассматривается среднее для данного интервала значение

$$f(x;t_1,t_2)=\frac{1}{t_2-t_1}\int_{t_1}^{t} f(x,t)\,dt \tag{22}$$

или

$$f(x;z_1,z_2)=\frac{1}{z_2-z_1}\int_{z_1}^{z_2} f(x,z-x)\,dz,$$

т. е. имеет место ограничение в среднем.

Теоретически можно представить себе такое поколение, которое характеризуется теми или иными демографическими показателями, описываемыми функцией $f(x;z_0)$ или $f(x;z_1,z_2)$, но в реальной действительности такого поколения не существует. Такую искусственную популяцию, в которой повозрастные демографические показатели равны повозрастным показателям, наблюдающимся в реальном населении в один и тот же момент или период времени, т. е. в которой $f(x)$ равно $f(x;z_0)$ или $f(x;z_1,z_2)$, называют, соответственно гипотетическим поколением момента z_0 или периода z_1, z_2.

На первый взгляд гипотетическое поколение, если не считать искусственного метода его построения, во всем напоминает реальное поколение, по образу и подобию которого оно создано. Однако на самом деле не так. Существует важное различие, которое заключается в следующем.

В реальном поколении значение любой демометрической функции в любом возрасте x отражает действительную историю данного поколения, что обеспечивает непротиворечивую связь между всеми функциями. В частности, если, как мы условились выше, говорить о смертности, то не может быть никакого противоречия, скажем, между функциями $\mu(x)$ и $l(x)$: если сила смертности в возрасте от 0 до 4 лет мала, то число доживающих до 5 лет обязательно отразит эту низкую смертность в предшествующих возрастах. Иначе говоря, если мы ограничим функцию $\mu(x,t)$ на прямую $t=t_0$ и исходя из полученной функции $\mu(x,t_0)$ формально построим функцию $l_{t_0}(x)$, а затем непосредственно ограничим $l(x,t)$ на прямую и получим $l(x,t_0)$, то всегда

$$l_{t_0}(x)=l(x,t_0), \tag{23}$$

т. е. вероятность дожития в реальном поколении всегда равна ограничению функции дожития на прямую $t=t_0$.

Иначе обстоит дело в гипотетическом поколении.

Здесь значения биометрических функций в каждом возрасте отражают историю своего поколения (реального). Применительно к смертности это значит, что в условиях меняющейся, особенно колеблющейся, смертности неизбежно возникновение неувязок и противоречий между различными функциями. В гипотетическом поколении сила смертности в возрасте 0—4 года может быть мала, но число доживших до 5 лет $l(5)$ может отражать высокую смертность (например, в результате эпидемии, которая вспыхнула 4 года назад) в возрасте 0—1 год в том поколении, которому к моменту наблюдения исполнилось 5 лет, и потому быть гораздо меньшим, чем это можно было ожидать исходя из соответствующих значений $\mu(x, z-x)$ при $0 \leqslant x \leqslant 5$ гипотетического поколения. В общем виде это значит, что если мы ограничим функцию $\mu(x, t)$ на изохрону z_0 и исходя из полученной функции $\mu(x, z_0)$ построим функцию $l_{z_0}(x)$, а затем непосредственно ограничим $l(x, t)$ на изохрону z_0, то в условиях меняющейся смертности получим

$$l_{z_0}(x) \neq l(x; z_0),$$

т. е. вероятность дожития в гипотетическом поколении не равна ограничению функции $l(x, t)$ на прямую z_0.

Сказанное справедливо для любых функций, описывающих смертность, если на их базе можно строить гипотетическое поколение. Ограничения их на вертикальную линию (или полосу) обязательно связаны каноническими соотношениями, записанными в формулах (6)—(19), а ограничения на изохрону или полосу изохрон могут не удовлетворять уравнениям (6)—(19), т. е. в условиях меняющейся смертности гипотетические поколения с разной исходной функцией могут не соответствовать друг другу. Поэтому гипотетическое поколение можно строить на базе какого-нибудь одного показателя, ибо два показателя могут оказаться взаимоисключающими.

Гипотетическое поколение может быть построено на базе такого показателя, из которого разворачиваются все другие. Формально имеется три возможности выбора такого показателя:

1) исходный показатель $\mu(x)$, приближенный с помощью ${}_v m(x)$ или ${}_v q(x)$;

2) исходный показатель $l(x)$;

3) исходный показатель $-l'(x)$, приближенный с помощью ${}_v d(x)$.

Несколько поясним, почему список возможных вариантов мы ограничили этими тремя. Прежде всего следует различать возможность расчета таблицы на базе данного показателя с той или иной степенью приближения и возможность развернуть систему функций из данной функции. Так, зная ${}_1q(x)$ для всех целых возрастов, можно построить таблицы смертности (однако ${}_1L(x)$ при этом будет найдено приближенно), но, даже зная ${}_1q(x)$ для всех x, мы не найдем точно ${}_{1/2}q(x)$, а зная ${}_{1/2}\ q(x)$, не найдем ${}_{1/4}\,q(x)$ и т. д.

Возможность строго вычислить все функции дает нам лишь *мгновенная* исходня функция. Именно все такие функции мы и собрали в нашем списке. Естественно, можно «сконструировать» иные функции с теми же свойствами, но мы ограничились функциями, так или иначе принятыми в демографии.

Поясним подробнее вариант 3. По определению производной,

$$-l'(x) = \frac{l(x) - l(x+\Delta x)}{\Delta x} + o(\Delta x), \qquad (24)$$

где $o(\Delta x) \to 0$ при $\Delta x \to 0$.

Так как $l(x) - l(x+\Delta x)$ есть вероятность для новорожденного умереть в интервале возрастов x, $x+\Delta x$, то показатель $-l'(x)$ — это аналог ${}_v d(x)$ на бесконечно малом интервале (т. е. когда $v \to 0$) в том же смысле, что $\mu(x)$ — аналог ${}_v q(x)$.

Обозначим $-l'(x)$ через $\lambda(x)$. Из формулы (24) следует, что

$$l(x) = 1 - \int\limits_0^x \lambda(y)\,dy. \qquad (25)$$

Исходя из любого из перечисленных показателей можно построить гипотетическое поколение, причем в зависимости от того, какой исходный показатель будет выбран, мы получим μ-гипотетическое, l-гипотетическое или λ-гипотетическое поколения. Хотя все они отражают одну и ту же реальную ситуацию, наблюдаемую в одном и том же населении в один и тот же момент, все эти гипотетические поколения будут обладать различными уровнями смертности. В табл. 1 приведены аналитические выражения для трех основных показателей μ, l и λ в трех типах гипотетических поколений. Из

Таблица 1

Сравнение основных показателей в гипотетических поколениях с различными исходными функциями

Тип гипотетического поколения	$\mu(x)$	$l(x)$	$\lambda(x)$
μ-гипотетическое поколение	$\mu(x, z-x)$	$l(x, z-x) \times \times e^{-\iint_{\Delta ABC} \mu'_t(x,t)dxdt}$	$\lambda(x, z-x) \times \times e^{-\iint_{\Delta ABC} \mu'_t(x,t)dxdt}$
l-гипотетическое поколение	$\mu(x, z-x) - -\int_0^x \mu'_t(y, z-x)dy$	$l(x, z-x)$	$\lambda(x, z-x) - l(x, z-x) \times \times \int_0^x \mu'_t(y, z-x)dy$
λ-гипотетическое поколение	$\dfrac{\mu(x, z-x)\, l(x, z-x)}{l(x, z-x) + \iint_{\Delta ABC} \lambda'_t(x, t)dxdt}$	$l(x, z-x) + +\iint_{\Delta ABC} \lambda'_t(x, t)dxdt$	$\lambda(x, z-x)$

П р и м е ч а н и е. Треугольник ABC есть треугольник на демографической сетке (см. рис. 1) с вершинами $A=(z-x, 0)$, $B=(z, 0)$, $C=(z-x, x)$.

приведенных выражений ясно, что в условиях изменяющейся смертности, т. е. когда $\mu'_t(x,t)$ и $\lambda'_t(x,t)$ не равны 0, эти показатели могут быть различными.

Мы не будем приводить вывод всех формул табл. 1. Скажем лишь, что этот вывод во всех случаях однотипен и не представляет большой сложности[5].

---

[5] Величина $\lambda(x)$ в l-гипотетическом поколении

$$\lambda_z^l(x) = \frac{\partial}{\partial x} l_z^l(x) = \frac{\partial}{\partial x} l(x, z-x) =$$

$$= l(x, z-x)\,\mu(x, z-x) + l(x,z-x) \int_0^x \mu'_t(y,z-x)\,dy,$$

а
$$\mu^l_z(x) = -\frac{\frac{\partial}{\partial x} l(x, z-x)}{l(x, z-x)} = \mu(x, z-x) - \int_0^x \mu'_t(y, z-x)\,dy.$$

Тем самым все формулы второй строки доказаны. По определению,

$$l_z^\mu(x) = e^{-\int_0^x \mu(y,\, z-y)dy},$$

$$\mu(x, z-x) = \mu^l_z(x) + \int_0^x \mu'_t(y, z-x)\,dy$$

и

$$\int_0^x \mu(\zeta, z-\zeta)\,d\zeta = \int_0^x \mu^l_z(\zeta)\,d\zeta + \int_0^x d\zeta \int_0^\zeta \mu'_t(y, z-\zeta)\,dy;$$

во втором интеграле сделаем подстановку $\zeta = z - t$, получим

$$\mu(x, z-x) = \int_0^x \mu^l_z(\zeta)\,d\zeta - \int_x^{z-x} dt \int_0^{z-t} \mu'_t(y,t)\,dy =$$

$$= \int_0^x \mu^l_z(\zeta)\,d\zeta + \iint_{\triangle ABC} \mu'_t(x,t)\,dx\,dt \qquad \text{(см. рис. 1).}$$

Итак,

$$l_z^\mu(x) = l_z^l(x)\, e^{-\iint_{\triangle ABC} \mu'_t(x,\, t)dxdt}.$$

Наконец,

$$\lambda_z^\mu(x) = -\frac{\partial}{\partial x} l_z^\mu(x) = \mu_z^\mu(x)\, l_z^\mu(x) = \mu(x, z-x) \times$$

$$\times\, l(x, z-x)\, e^{-\iint_{\Delta ABC} \mu'_t(x,\, t)dxdt} = \lambda(x, z-x)\, e^{-\iint_{\Delta ABC} \mu'_t(x,\, t)dxdt}.$$

Таким образом, мы доказали и первую строку табл. 1. Третья строка доказывается аналогично.

Приведенные в табл. 1 формулы позволяют установить определенные зависимости между одними и теми же показателями для различных гипотетических поколений. Так, если имеется направленное изменение смертности, причем этот процесс идет монотонно, т. е. на протяжении рассматриваемого периода времени μ'_t и λ'_t меньше 0, то, как видно из табл. 1,

$$l\,(x, z-x) < l^{\mu}_{z}\,(x) < l^{\lambda}_{z}\,(x); \qquad (26)$$

$$\mu^{l}_{z}\,(x) > \mu\,(x, z-x) > \mu^{\lambda}_{z}\,(x); \qquad (27)$$

$$\lambda^{l}_{z}\,(x) > \lambda^{\mu}_{z}\,(x) > \lambda\,(x, z-x)\,. \qquad (28)$$

Таким образом, в любом возрасте x, до которого наблюдается монотонное снижение смертности, наибольшие значения чисел доживающих, наименьшая скорость снижения этих чисел и наименьшая сила смертности оказываются в λ-гипотетическом поколении, показатели которого можно поэтому считать наиболее оптимистическими. Этим объясняется, в частности, то, что введенный В. Я. Буняковским метод построения таблиц смертности с исходными $_{v}d(x)$ (подробное изложение метода см. в работе: «Курс демографии» под ред. А. Я. Боярского, с. 156—157), аналогом которых, как уже отмечалось, является функция $\lambda(x)$ дает самую низкую смертность. Важно отметить, что метод В. Я. Буняковского дает иное гипотетическое поколение, нежели метод с исходными $\mu(x)$. Мы нисколько не умаляем значения этого метода, однако следует признать, что сочетание в одной таблице двух типов гипотетического поколения, как это происходит при применении метода В. Я. Буняковского для детских возрастов, и так называемого демографического метода с исходным показателем q_x или m_x для прочих возрастов, не является целесообразным.

Обычно, говоря о гипотетическом поколении в связи с изучением смертности, имеют в виду μ-гипотетическое поколение. Ему соответствует большинство современных таблиц смертности, в частности все таблицы, построенные так называемым демографическим методом.

Предпочтение, которое отдается μ как исходному показателю при построении гипотетического поколения, не случайно. Сила смертности есть мера интенсивности смертности в различных возрастах, причем нет никакой обязательной зависимости между отдельными значениями этого показателя; поэтому ряд $\mu(x)$ не может быть внутренне противоречивым. Хотя, по общему правилу, абсолютные значения $\mu(x)$ монотонно нарастают (если не считать их уменьшения в связи с парадоксом детской смертности), немонотонность функции $\mu(x)$ не является теоретически или практически невозможной.

Если же при построении гипотетического поколения исходить из $l(x)$ или $\lambda(x)$, то легко можно прийти к абсурду: функция $l(x)$ в условиях сильно меняющейся смертности может оказаться немонотонной (что бессмысленно, так как число доживающих до возраста x не может быть меньше числа доживающих до возраста

e '$(1+x\int\limits_{\infty}^{0}\lambda(x)dx$, в принципе равный общему числу умерших, т. е. 1, может быть от единицы отличен.

Как мы уже отмечали выше, в условиях меняющейся смертности смертность гипотетического[6] поколения не соответствует в целом смертности никакого реального поколения. Из табл. 1 следует связь между функциями $l(x)$ в реальном и гипотетическом поколениях, сходящимися в одной точке C (см. рис. 1):

$$l(x, z-x) = l_z(x)\,e^{-\iint\limits_{\Delta ABC} \mu'_t(x,\,t)\,dx\,dt} \qquad (29)$$

Полезно проиллюстрировать эти формулы и различия на практическом примере. На рис. 2 приведены погодные вероятности умереть мужчин в Италии (детские возрасты) в реальных поколениях 1950—1956 гг.[7].

Имеющаяся в таблице информация позволяет построить таблицы смертности как для реального поколения 1950—1951 гг., так и для гипотетического поколения 1955—1956 гг. в смысле μ^-, λ^-, l^- гипотетических поколений. Так, $l(6)$ в реальном поколении, равная $l(6)$ в l-гипотетическом поколении, равна 0,9108244, в μ-гипотетическом поколении 0,936756, в λ-гипотетическом поко-

---

[6] Всюду в дальнейшем под гипотетическим поколением мы будем понимать μ-гипотетическое поколение.

[7] A. Naddeo. La mortalita in Italia dopo il 1950. Instituto di demografia Facolta'di scienze statistiche demografiche e attuariali dell'universita'di Roma. Roma, 1959.

лении, рассчитанном по методу В. Я. Буняковского — 0,936845.

Отношение

$$\frac{l(6, t)}{l_z^{\mu}(6)} = 0{,}9723.$$

Мы попытались приблизительно оценить $\iint\limits_{\triangle ABC} \mu^1(x, t)\,dx\,dt$ (см. рис. 2), пренебрегая тем, что реальные и гипотетические поколения интервальные, а не мгновенные. При этом мы считаем функцию $\mu(x, t)$ постоянной по x внутригодичных интервалов и линейно

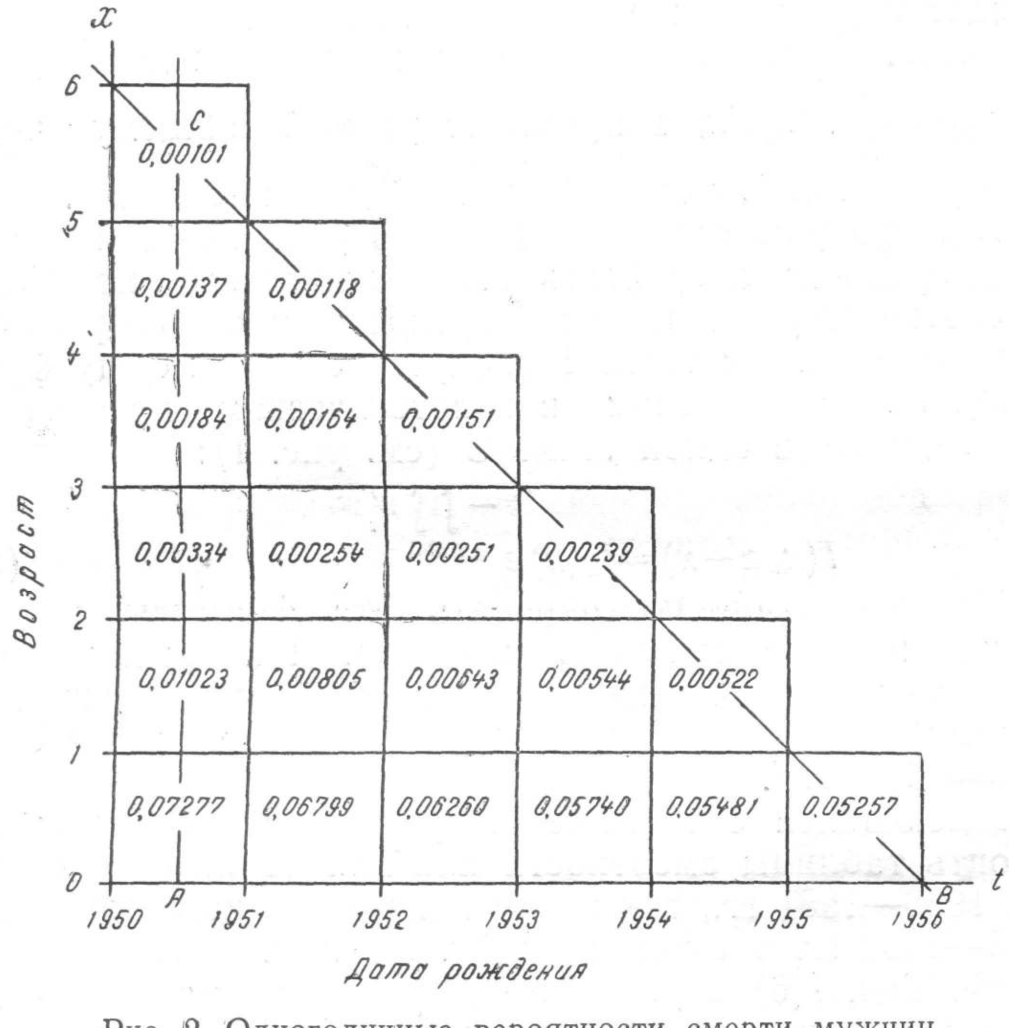

Рис. 2. Одногодичные вероятности смерти мужчин, родившихся в 1950—1956 гг. (${}_1q(x)$), Италия.

интерполировали ее по t. Полученный результат — 0,0289. Величина $e^{-0{,}0289} = 0{,}9715$. Некоторое расхождение вызвано погрешностями интегрирования.

Сравнение функций дожития в реальном и гипотетическом поколениях может быть продолжено и применительно к интервальным показателям.

Аналогично формуле (29) о связи $l_z(x)$ и $l(x, z-x)$ доказываются следующие соотношения (30). Положим $p(x_1, x_2) = {}_{x-x_1}p(x_1)$. Пусть $t_1 = z - x_2$, $t_2 = z - x_1$ (см. рис. 3). Имеют место следующие равенства:

$$\frac{p_z(x_1, x_2)}{p(x_1, x_2; t_1)} = e^{\iint\limits_{\Delta EDG} \mu'_t(x, t)\,dx\,dt};$$

$$\frac{p(x_1, x_2; t_2)}{p_z(x_1\ x_2)} = e^{\iint\limits_{\Delta GEF} \mu'_t(x, t)\,dx\,dt} \qquad (30)$$

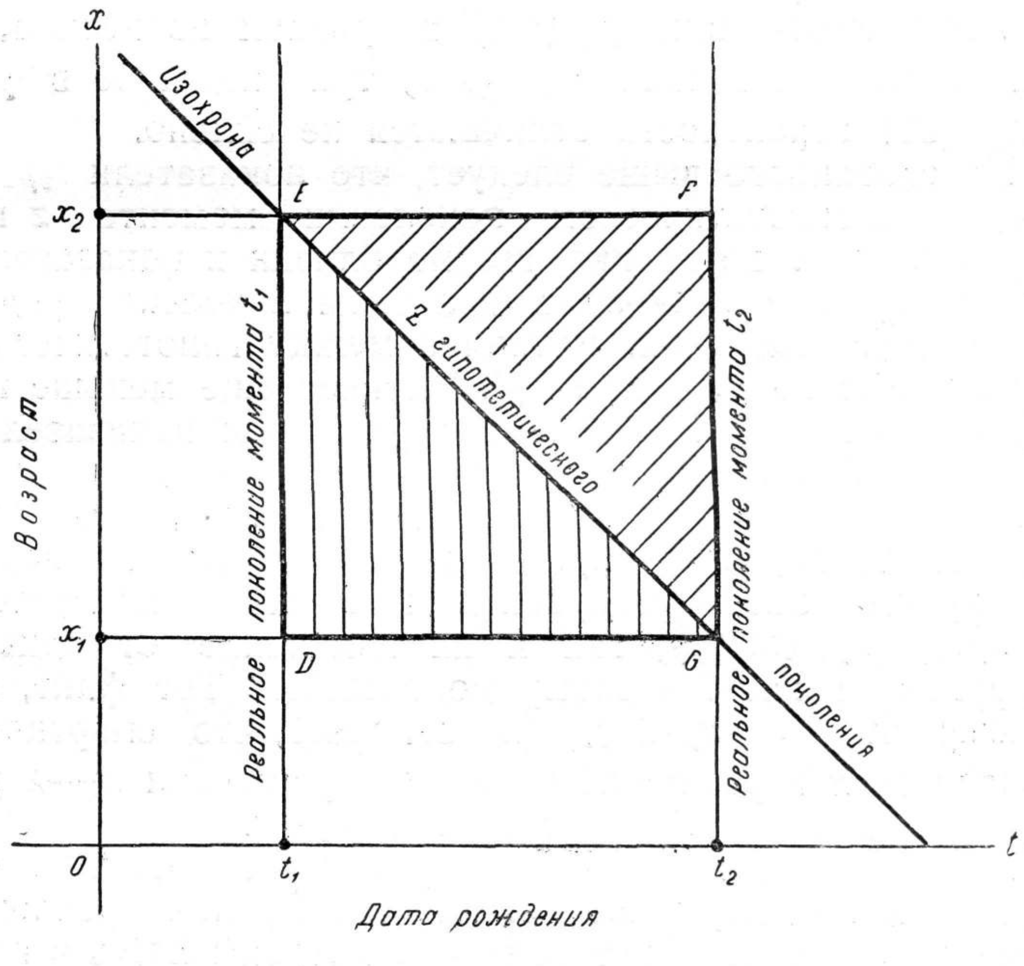

Рис. 3.

Если допустить, что $\mu'_t < 0$ в четырехугольнике $EDGF$, то легко убедиться, что

$$p(x_1, x_2, t_1) < p_z(x_1\, x_2) < p(x_1, x_2, t), \qquad (31)$$

что доказывает, что смертность сегодняшнего μ-гипотетического поколения ближе к смертности будущих поколений, чем смертность прошлых поколений, в условиях направленного изменения смертности. Отметим, что в случае l- или λ-гипотетического поколения аналогичное утверждение не верно.

Формулы (30) характеризуют связь этих вероятностей.

Стоящие в показателе величины $\iint\limits_{\triangle EDG} \mu'_t\,(x, t)dxdt$ $\iint\limits_{\triangle GEF} \mu'_t\,(x, t)dxdt$ есть средние значения изменения силы смертности на этих участках демографической плоскости, помноженные на площадь соответствующих треугольников.

Так как величина $\mu'_t\,(x, t)$ в среднем не велика, то при небольшой величине $\mu'_t\,(x, t)$ три входящие в формулу (31) вероятности отличаются не сильно.

Из сказанного выше следует, что показатели ${}_vp_z(x)$ и ${}_vq_z(x)$ в гипотетическом поколении момента z при условии, что v не велико, весьма близки к показателям ${}_vp(x, z—x)$ и ${}_vq(x, z—x)$ в реальном поколении родившихся в момент $z—x$. Различие между гипотетическим и реальным показателем ${}_vq(x)$ станет еще меньше при переходе от моментных к интервальным показателям. Это позволяет строить таблицы смертности для гипотетического поколения, принимая, например, ${}_vq_{z_1 z_2}\,(x)$ равным ${}_vq(x; z_1—x, z_2—x)$. Эти же соображения позволяют просто интерпретировать как сами показатели ${}_vq_z(x), {}_vp_z(x), {}_vm_z(x)$, так и их изменение от одного к другому гипотетическому поколению. Тот факт, например, что ${}_vq_z\,(x) < {}_vq_{z_1}\,(x)$ означает, что смертность в возрасте x в реальном поколении рождения $z_2—x$ ниже, чем в поколении рождения $z_1—x$.

Прочие характеристики смертности гипотетического поколения не допускают столь конкретной реальной интерпретации, но могут быть рассмотрены лишь в рамках первоначальной модели поколения, родившегося и прожившего всю жизнь в условиях смертности сегодняшнего дня. Вопрос о том, в какой степени эти показатели характеризуют сегодняшний уровень смертности и как их динамика от одного гипотетического поколения к другому отражает реальное изменение смертности, до сих пор остается открытым. Наконец, весьма важной является задача найти один или несколько показателей, количественно оценивших такое понятие, как уровень смертности, и даже отвечающих на вопрос о благоприятности или неблагоприятности или, точнее, о степени благоприятности изменения смертности.

Оценка этих изменений не вызывает трудностей, когда $_vq(x)$ или $\mu(x)$ снижается с ростом t во *всех* возрастах. Сложность поставленной проблемы состоит в том, что такое снижение смертности не единственно возможное, ибо часто наряду со снижением смертности в одних возрастах наблюдается ее рост в некоторых других. Кроме того, темп снижения смертности в разных возрастах различен, а различные сочетания повозрастных темпов снижения дают и разные итоговые уровни смертности.

3. Анализ возможных универсальных характеристик уровня смертности

Существует ряд показателей, претендующих на роль измерителей уровня смертности. Они представляют собой те или иные средние (типа средний возраст умерших).

Они могут быть вычислены как в гипотетическом поколении любого типа, или то же самое, в стационарном населении с данным уровнем смертности, так и в реальном поколении и во всем населении в данный период времени.

В разных случаях эти средние разнятся количественно и несут разную информацию.

Первый такой показатель — средняя продолжительность предстоящей жизни в возрасте x лет:

$$e(x) = \frac{\int\limits_x^\infty l(y)\,dy}{l(x)}\,. \tag{32}$$

В реальном поколении величина $e(x;\ t_1,\ t_2)$ действительно равна среднему числу лет предстоящей жизни у лиц, доживших до возраста x лет. Известно, что

$$e(x)=\frac{\int\limits_x^\infty (y-x)\lambda(y)\,dy}{l(x)},$$

т. е. средняя продолжительность предстоящей жизни $e(x)$ равна среднему возрасту умерших в возрасте свыше x, уменьшенному на x. Величина $e(x)$ может быть вычислена и в гипотетическом поколении:

$$e(x)=\frac{\int\limits_x^\infty l_z(y)\,dy}{l_z(x)}.$$

Аналогично может быть вычислен средний возраст живущих в возрасте свыше x в стационарном населении с данным уровнем смертности

$$g(x)=x+\frac{\int\limits_x^\infty (y-x)\,l(y)\,dy}{\int\limits_x^\infty l(y)\,dy}. \tag{33}$$

На наш взгляд, показатель $g(x)$ заслуживает гораздо больше внимания, чем ему уделяется в настоящее время.

Рассмотрим следующую задачу. Допустим, что мы строим гипотетическое поколение периода z_1, z_2, на базе повозрастных чисел умерших и повозрастных численностей населения.

По одному гипотетическому поколению судить об изменении смертности нельзя; поэтому пусть у нас имеется и информация о числах рождений за предшествующие годы. Тогда мы можем построить не только μ-гипотетическое поколение, но и l-гипотетическое поколение и λ-гипотетическое. Имея в наличии три гипотетических поколения данного периода или данного момента, уже можно оценить тенденции изменения смертности, опираясь на формулы, приведенные в табл. 1. Так, если

$$l(x, z-x)<\overset{\mu}{l_z}(x)$$

или (34)

$$\lambda(x, z-x)<\overset{\mu}{\lambda_z}(x),$$

то это означает, что в треугольнике ABC (см. рис. 1), точнее, в поколениях родившихся на отрезке $[z-x, x]$ за годы, предшествующие z, наблюдалась общая тенденция снижения смертности. Точно так же факт, что

$$e^{\mu}_{z}(0) > e^{l}_{z}(0), \tag{35}$$

говорит об общей тенденции снижения смертности в поколениях, доживших до момента z. На практике величину $e^{\mu}_{z}(0)$ можно вычислить как среднюю продолжительность предстоящей жизни, а величину $e^{l}_{z}(0)$, зная числа родившихся $N(t, t+1)$ и числа живущих в

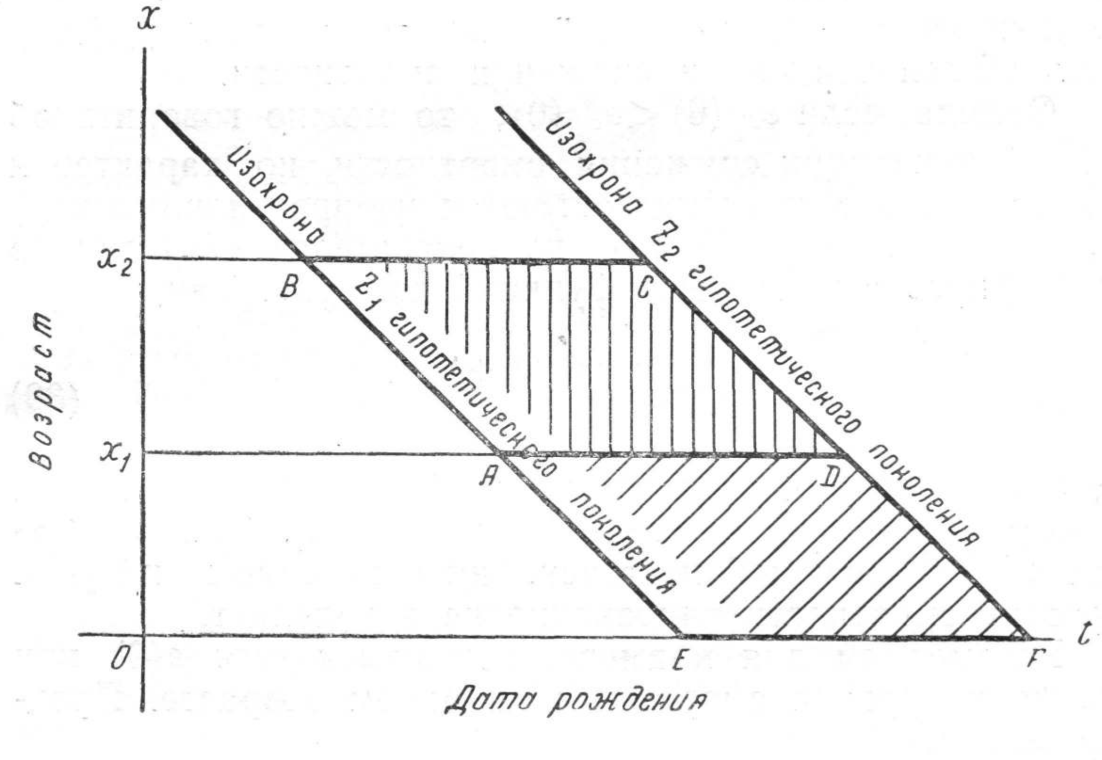

Рис. 4.

момент z в интервале возрастов $(x, x+1) - \hat{S}(z,x)$, найти по приближенной формуле

$$e^{l}_{z}(0) \approx \sum_{n=0}^{\omega} \frac{\hat{S}(z,n)}{N(z-n-1,\ z-n)}, \tag{36}$$

где ω — предельный возраст живущих.

Допустим теперь, что мы сравниваем два гипотетических поколения моментов времени z_1 и z_2 (см. рис. 4), Точно так же, как выводятся формулы (30), можно показать, что

$$\frac{p_{z_1}(x_1, x_2)}{p_{z_2}(x_1 x_2)} = e^{\iint\limits_{ABCD} \mu'_t(x, t)dxdt} \tag{37}$$

или

$$\frac{l_{z_1}(x_1)}{l_{z_2}(x_1)} = e^{\iint\limits_{ADEF} \mu'_t(x, t)dxdt}$$

Ясно, что если

$$p_{z_1}(x_1, x_2) < p_{z_2}(x_1, x_2), \tag{38}$$

то можно утверждать, что в поколениях, пронизывающих четырехугольник ABCD в возрастах $x_1 \leqslant x \geqslant x_2$, есть общая тенденция снижения смертности.

Отсюда, если $e_{z_1}(0) < e_{z_2}(0)$, то можно говорить об общей тенденции снижения смертности, но характер и качество этого снижения остаются неопределенными.

Легко проверить, что из показателя $e(x)$ можно развернуть всю систему функций дожития. Так,

$$\mu(x) = \frac{e'_x(x) + 1}{e(x)} \tag{39}$$

и, следовательно, по значениям функции $e(x)$ во всех точках определяется уровень смертности[8]. Однако знание $e(x)$ ни в одной из точек, ни в конечном их ряде уровень смертности однозначно не определяет.

Рассмотрим для примера, как изменяется $e(0)$ при снижении уровня смертности в одном возрасте. Пред-

---

[8] Действительно,

$$e'_x(x) = \frac{d}{dx}\,\frac{\int\limits_x^\infty l(y)dy}{l(x)} =$$

$$= \frac{-l(x)l(x) - [-\mu(x)l(x)]\int\limits_x^\infty l(y)dy}{l^2(x)};$$

отсюда $e'_x(x) = -1 + \mu(x)e(x)$; из этого равенства непосредственно следует (39).

положим, что уровень смертности изменился на отрезке x, $x+v$ так, что ${}_vd(x)$ уменьшилось на $\Delta_{d,v}L(x)$ увеличилось на Δ_L, а вне этого отрезка никаких изменений не произошло.

Величина $e(0)$ до изменения смертности равна $\int\limits_0^x l(y)\,dy + {}_vL(x) + l(x+v)\,e(x+v)$, а после изменения равна $\tilde{e}(0) = \int\limits_0^x l(y)\,dy + {}_vL(x) + (l(x+v) + \Delta_d)\,e\,(x+v) + \Delta_L$. Ясно, что

$$\tilde{e}(0) - e(0) = \Delta_L + \Delta_d(e(x+v)).$$

Если снижение смертности равномерно на отрезке $[x, x+v]$, то

$$\tilde{e}(0) - e(0) = \Delta_d\left(\frac{v}{2} + e(x+v)\right)\cdot \tag{40}$$

Отсюда следует, что снижение смертности в возрасте x тем сильнее влияет на $e(0)$, чем больше $e(x+v)$, т. е. снижение смертности в молодых возрастах существеннее, чем изменение в старческих.

Исходя из равенства (40) можно записать более общую формулу изменения $e(0)$, если в возрастах $x_1, x_2, \ldots, x_n$ (т. е. на интервалах x_i, x_i+1 при $i=1, \ldots, n$) смертность равномерно снизилась на Δ_i; однако такая формула вряд ли прояснит дело.

Показатель, адекватно характеризующий смертность, в принципе реален, ибо в целом ряде работ[9], связанных с модельными таблицами смертности, если не доказано, то убедительно показано, что множество функций $l(x)$, описывающих смертность человеческих популяций, не бесконечномерно, а имеет размерность порядка четырех.

---

[9] См., например: *S. Lederman, J. Breas.* Les dimensions de la mortalité. — «Population», vol. 14, 1959, № 4; «Factor Analysis of Sex — Age Specific Death Rates». — «Population Bulletin of the United Nations», 1962, № 6, p. 149—201.