

长期护理保险制度中的失能群体结构性变化

——基于某副省级城市的试点证据

吴舒祥 臧文斌 叶小兰 赵绍阳

目录

附录 I 长护险试点初期的三类失能群体	1
附录 II 样本城市的长护险渐进式扩展过程	3
附录 III 正文(3)式证明	4
附录 IV 状态转移数据提取过程、模型拟合方式与代码逻辑	6
附录 V 附表及附图	9

附录 I 长护险试点初期的三类失能群体

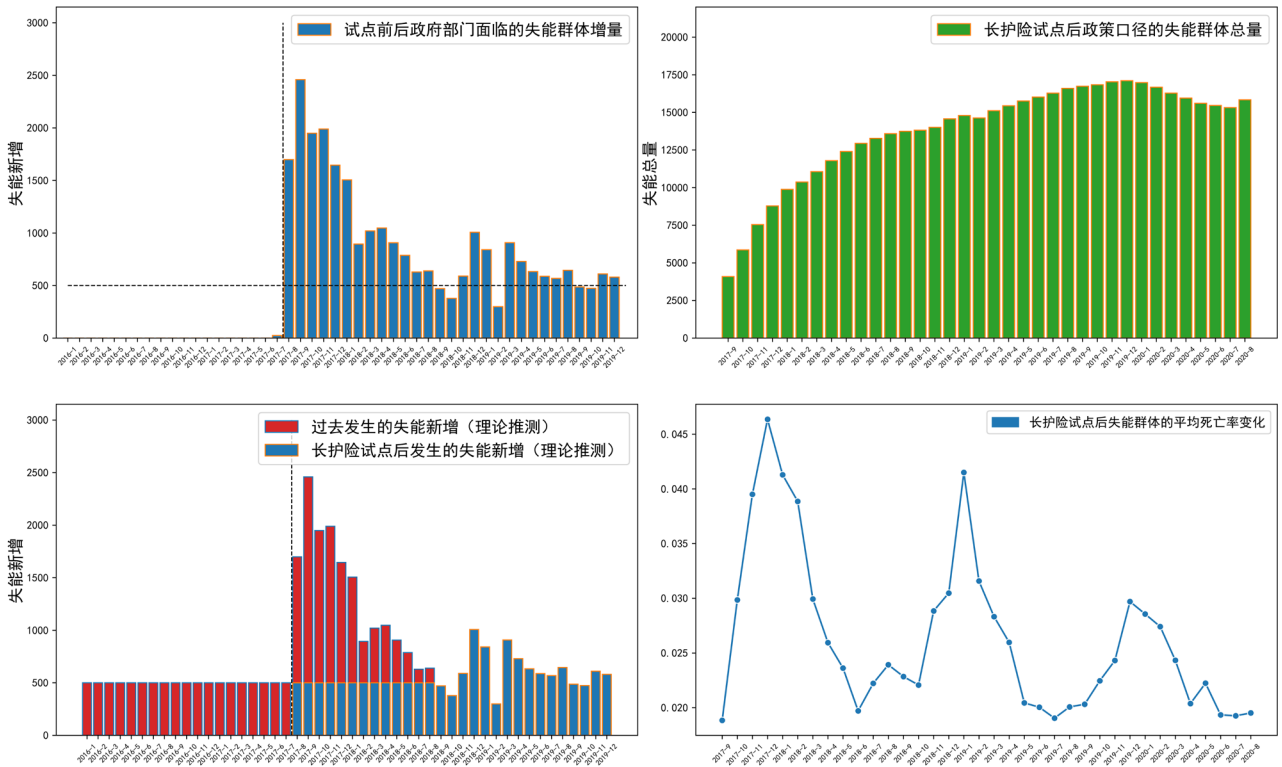


图 11 某副省级城市长护险试点前后的失能群体数据变化

如图 11 所示，第一类失能群体即长护险制度实施前失能且死亡的群体，这类群体从失能 \Rightarrow 死亡都未获得长护险支持，也不会产生长护险基金支出；第二类群体即长护险制度实施前失能、长护险制度实施后死亡的群体，也是长护险试点初期获得长护险支持的主要群体，也表现为某地区首次启动长护险制度时的失能人口。这类群体仅能在“失能 \Rightarrow 死亡”部分期间产生长护险基金支出；第三类失能群体即长护险制度实施后失能且死亡的群体，在“失能 \Rightarrow 死亡”全部期间产生长护险基金支出。本文认为在同等条件下，即使考虑失能余寿波动，同样人数的第三类失能群体也会比同样人数的第二类失能群体产生更多的长护险待遇支付次数。同时，第三类失能群体由于能获得更长平均时间的长护险支持，长护险照护支持下失能人员可能发生健康恶化速度减缓，还可能加剧两类群体的平均健康水平差异（如附录图 A3 绿色部分所示）。

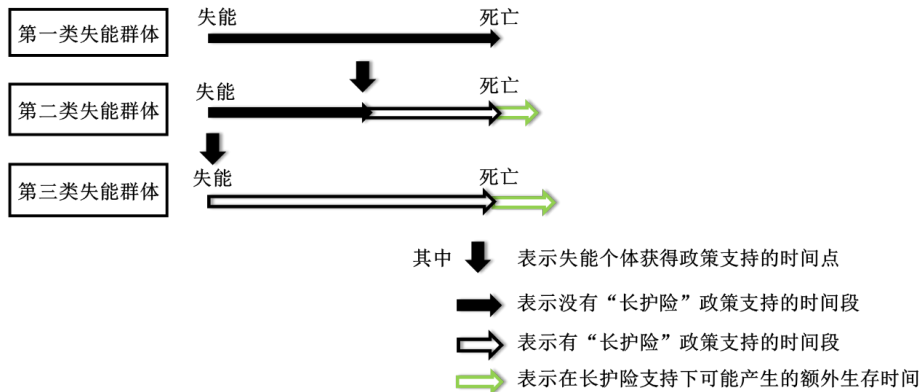


图 12 长护险制度推行初期的三类失能群体

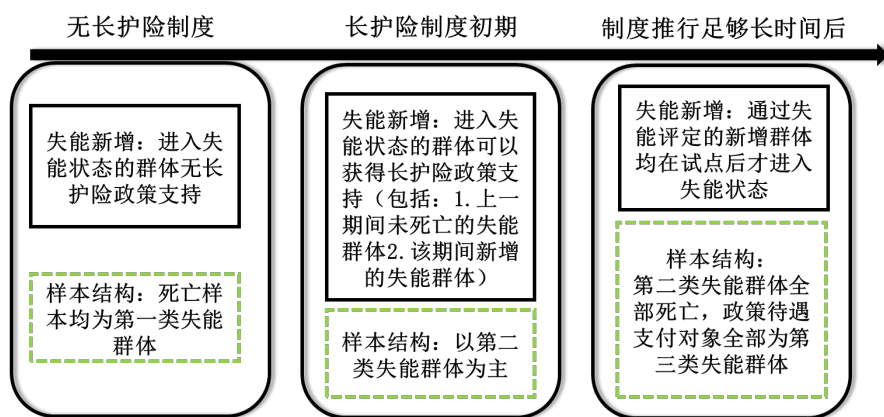


图 13 长护险试点前后 (或待遇支付范围变化前后) 的失能样本结构变化

附录 II 样本城市的长护险渐进式扩展过程

表 111 样本城市长护险政策的具体安排

序号	时间	描述	参保范围	保障范围
时间点 1	2017 年 7 月 1 日	首轮试点启动	城镇职工基本医疗保险 参保人员	重度失能群体（一级、二级、三 级）
时间点 2	2018 年 12 月 1 日	保障范围扩大	城镇职工基本医疗保险 参保人员	重度失能群体（一级、二级、三 级）+失智群体
时间点 3	2021 年 1 月 1 日	参保范围扩大	城镇职工+城乡居民+省 本级职工	重度失能群体（一级、二级、三 级）+失智群体
时间点 4	2022 年 7 月 1 日	参保范围扩大+ 保障范围扩大+待 遇提高	城镇职工+城乡居民+省 本级职工+学生儿童	重度失能群体（一级、二级、三 级）+ +中度失能（二级、三级）+失智群 体
时间点 5	2023 年 4 月 1 日	待遇提高	城镇职工+城乡居民+省 本级职工+学生儿童	重度失能群体（一级、二级、三 级）+ +中度失能（二级、三级）+失智群 体

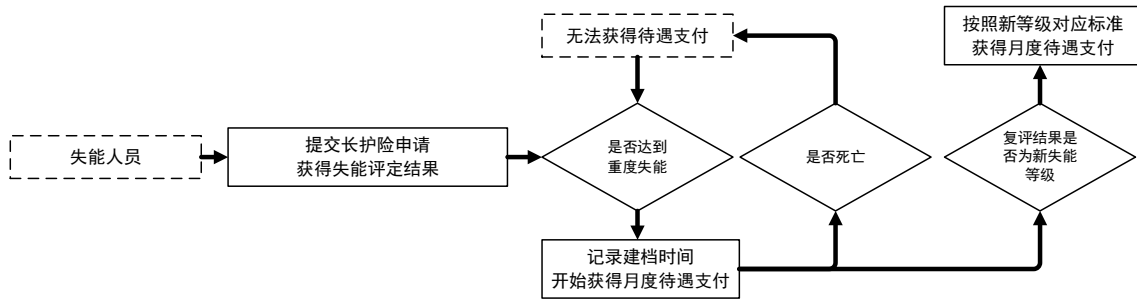


图 111 长护险试点数据统计过程

图 II2 报告了样本城市在经济发展水平、人口结构和医疗资源与其他城市的比较结果。其中，经济发展水平由“人均 GDP”数据进行量化，来自 2019 年度中国城市统计年鉴；医疗资源由每“千人口床位数”数据进行量化，来自 2019 年度中国城市统计年鉴；人口结构数据由“65 岁及以上人口比重”数据进行量化，来自第七次全国人口普查数据。由于长护险 15 个首批试点城市均为经济发展和医疗资源水平相对较高的城市，样本城市的人均 GDP 和每千人口床位数均高于全国平均水平。此外，样本城市 65 岁以上人口占比处于全国平均水平，其人口结构具有较高的全国代表性。

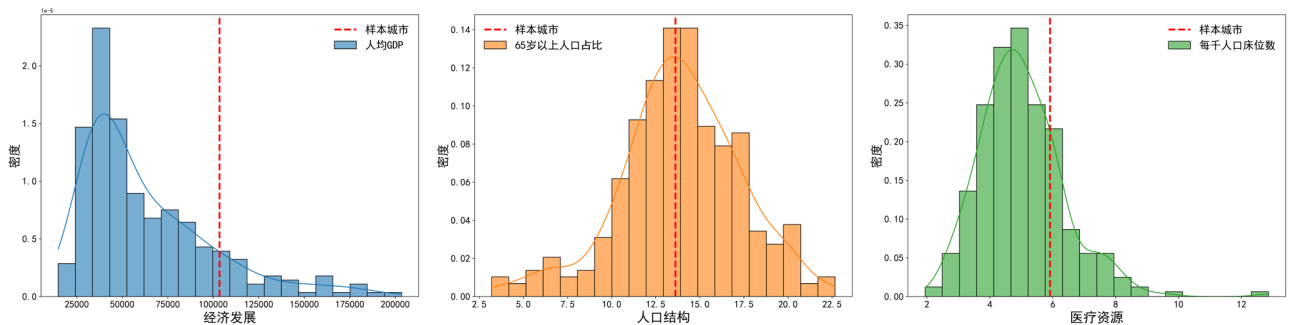


图 112 样本城市的全国代表性

附录 III 正文式 (3) 证明

证明:

$$\mathbf{P}_{ij}(t, \Delta t_1 + \Delta t_2)_{(i,j \in \mathbf{I})} = \mathbf{P}_{ij}(t, \Delta t_1) \cdot \mathbf{P}_{ij}(t + \Delta t_1, \Delta t_2)_{(i,j \in \mathbf{I})} \quad \text{正文式 (3)}$$

证:

由于 $p_{ij}(t, \Delta t_1 + \Delta t_2)_{(i,j \in \mathbf{I})}$ 的含义为一个年龄为 t 、健康状态为 i 的参保人在 $\Delta t_1 + \Delta t_2$ 之后健康状态转移到 j 的概率, 仅限定了 t 时刻状态和 $t + \Delta t_1 + \Delta t_2$ 时刻状态, 因此在 $t + \Delta t_1$ 时刻可以为状态空间 \mathbf{I} 中任意一个非吸收状态 k (如式 (III 1) 所示)。因此, 如果需要将 $p_{ij}(t, \Delta t_1 + \Delta t_2)_{(i,j \in \mathbf{I})}$ 由两个阶段的状态转移概率表示, 即附录式 (III 2)。

$$\left\{ \begin{array}{l} t \Rightarrow t + \Delta t_1 \Rightarrow t + \Delta t_1 + \Delta t_2 \\ i \Rightarrow \forall k \in \mathbf{I} \Rightarrow j \end{array} \right. \quad \text{(III 1)}$$

$$\begin{aligned} p_{ij}(t, \Delta t_1 + \Delta t_2)_{(i,j \in \mathbf{I})} &= p_{i,k=1}(t, \Delta t_1) \cdot p_{k=1,j}(t + \Delta t_1, \Delta t_2) + p_{i,k=2}(t, \Delta t_1) \cdot p_{k=2,j}(t + \Delta t_1, \Delta t_2) + \dots \\ &= \sum_{k=1} p_{ik}(t, \Delta t_1) \cdot p_{kj}(t + \Delta t_1, \Delta t_2)_{(i,j,k \in \mathbf{I})} \end{aligned} \quad \text{(III 2)}$$

根据正文式 (1) 关于状态转移矩阵 \mathbf{P} 的定义, 我们即可以将式 (III 2) 代入状态转移矩阵 $\mathbf{P}_{ij}(t, \Delta t_1 + \Delta t_2)$ 的矩阵表达式中, 即可得到正文式 (3) 等号左边的表达式, 即式 (III 3)。

$$\begin{aligned} \mathbf{P}_{ij}(t, \Delta t)_{(i,j \in \mathbf{I})} &= \begin{bmatrix} p_{11}(t, \Delta t) & p_{12}(t, \Delta t) & \dots & p_{1n}(t, \Delta t) \\ p_{21}(t, \Delta t) & p_{22}(t, \Delta t) & \dots & p_{2n}(t, \Delta t) \\ \dots & \dots & \dots & \dots \\ p_{n1}(t, \Delta t) & p_{n2}(t, \Delta t) & \dots & p_{nn}(t, \Delta t) \end{bmatrix} \quad \text{正文式 (2)} \\ \mathbf{P}_{ij}(t, \Delta t_1 + \Delta t_2)_{(i,j \in \mathbf{I})} &= \begin{bmatrix} p_{11}(t, \Delta t_1 + \Delta t_2) & p_{12}(t, \Delta t_1 + \Delta t_2) & \dots & p_{1n}(t, \Delta t) \\ \dots & \dots & \dots & \dots \\ p_{n1}(t, \Delta t_1 + \Delta t_2) & p_{n2}(t, \Delta t_1 + \Delta t_2) & \dots & p_{nn}(t, \Delta t_1 + \Delta t_2) \end{bmatrix} \\ &= \begin{bmatrix} \sum_{k=1} p_{1k}(t, \Delta t_1) \cdot p_{k1}(t + \Delta t_1, \Delta t_2) & \dots & \sum_{k=1} p_{1k}(t, \Delta t_1) \cdot p_{kn}(t + \Delta t_1, \Delta t_2) \\ \dots & \dots & \dots \\ \sum_{k=1} p_{nk}(t, \Delta t_1) \cdot p_{k1}(t + \Delta t_1, \Delta t_2) & \dots & \sum_{k=1} p_{nk}(t, \Delta t_1) \cdot p_{kn}(t + \Delta t_1, \Delta t_2) \end{bmatrix}_{n \times n} = \text{正文式 (3) (左)} \end{aligned}$$

同理, 正文式 (3) (右) = $\mathbf{P}_{ij}(t, \Delta t_1) \cdot \mathbf{P}_{ij}(t + \Delta t_1, \Delta t_2)$

$$\begin{aligned}
 &= \begin{bmatrix} p_{11}(t, \Delta t_1) & p_{12}(t, \Delta t_1) & \cdots & p_{1n}(t, \Delta t_1) \\ p_{21}(t, \Delta t_1) & p_{22}(t, \Delta t_1) & \cdots & p_{2n}(t, \Delta t_1) \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1}(t, \Delta t_1) & p_{n2}(t, \Delta t_1) & \cdots & p_{nn}(t, \Delta t_1) \end{bmatrix} \cdot \begin{bmatrix} p_{11}(t + \Delta t_1, \Delta t_2) & \cdots & p_{1n}(t + \Delta t_1, \Delta t_2) \\ p_{21}(t + \Delta t_1, \Delta t_2) & \cdots & p_{2n}(t + \Delta t_1, \Delta t_2) \\ \cdots & \cdots & \cdots \\ p_{n1}(t + \Delta t_1, \Delta t_2) & \cdots & p_{nn}(t + \Delta t_1, \Delta t_2) \end{bmatrix}_{n \times n} \\
 &= \begin{bmatrix} \sum_{k=1}^n p_{1k}(t, \Delta t_1) \cdot p_{k1}(t + \Delta t_1, \Delta t_2) & \cdots & \sum_{k=1}^n p_{1k}(t, \Delta t_1) \cdot p_{kn}(t + \Delta t_1, \Delta t_2) \\ \cdots & \cdots & \cdots \\ \sum_{k=1}^n p_{nk}(t, \Delta t_1) \cdot p_{k1}(t + \Delta t_1, \Delta t_2) & \cdots & \sum_{k=1}^n p_{nk}(t, \Delta t_1) \cdot p_{kn}(t + \Delta t_1, \Delta t_2) \end{bmatrix}_{n \times n} = \text{正文式 (3) (左)}
 \end{aligned}$$

证毕

附录 IV 状态转移数据提取过程、模型拟合方式与代码逻辑

每个失能参保人的健康状态转移变化主要由其死亡信息和复评信息构成。根据失能群体在截止日的健康状态，所有研究样本可分为死亡样本（12221 个）和非死亡样本（14221 个）。每个通过失能评定的样本 $(t_0, S(t_0))$ (t_0 代表年龄， $S(t_0)$ 代表初始评估等级) 的健康状态转移过程 $\{S(t)\}$ 主要分为三类。第一类为

死亡样本，其健康状态变化过程为 $(t_0, S(t_0)) \rightarrow (t_1, S(t_1)=4)$ ，即 $\begin{cases} S(t) = S(t_0) & , t_0 \leq t < t_1 \\ S(t) = 4 & , t \geq t_1 \end{cases}$ ；第二类为失能评

估等级出现变化的样本，其健康状态变化过程为 $(t_0, S(t_0)) \rightarrow (t_1, S(t_1))$ ，即 $\begin{cases} S(t) = S(t_0) & , t_0 \leq t < t_1 \\ S(t) = S(t_1) & , t \geq t_1 \end{cases}$ ；第三

类为到截止日期既未死亡，失能等级也未变化的样本，其健康状态变化过程为 $(t_0, S(t_0)) \rightarrow (t_{\max}, S(t_0))$ ，即 $S(t)=S(t_0)$ ($t_0 \leq t \leq t_{\max}$)。在该附录内容中，本文以两个真实失能样本为例，详细报告了状态转移数据的提取过程。

表 IV1 状态转移数据提取过程

<p>失能人员 1: 该失能人员 73 岁评定为重度失能 2 级，获得重度失能 2 级待遇支付 20 个月，复评为重度失能 3 级，获得重度 3 级待遇支付 16 个月死亡。</p>
<p style="text-align: center;">失能人员 1 的状态转移数据提取过程:</p> <p>(1) 首先我们统一时间单位，将 73 岁换算位月份，即 $73 \times 12=876$ 个月；</p> <p>(2) 其次，我们提取该样本的健康状态变化过程$\{S(t)\}$，如下式所示：</p> $\underbrace{\{S(876) = 2\}}_{S(t)=2} \Rightarrow \dots \Rightarrow \underbrace{\{S(896) = 2\} \Rightarrow \{S(896) = 3\}}_{S(t)=3} \Rightarrow \dots \Rightarrow \{S(1014) = 3\} \Rightarrow \{S(1014) = 4\}$
<p>失能人员 2: 该失能人员 70 岁评定为重度失能 1 级，获得重度 1 级待遇支付 37 个月后依然保持重度失能 1 级（在数据截至试点保持存活）。</p>
<p style="text-align: center;">失能人员 2 的状态转移数据提取过程:</p> <p>(1) 首先我们统一时间单位，将 70 岁换算位月份，即 $70 \times 12=840$ 个月；</p> <p>(2) 其次，我们提取该样本的健康状态变化过程$\{S(t)\}$，如下式所示：</p> $\underbrace{\{S(870) = 1\} \Rightarrow \{S(871) = 1\} \Rightarrow \dots \Rightarrow \{S(906) = 1\}}_{S(t)=1} \Rightarrow \{S(907) = 1\}$

在非时齐 Markov 状态转移过程中，给定初始状态 $S(t)=i$ 、转移间隔 Δt 和未来状态 $S(t+\Delta t)=j$ 时，状态转移概率 $p_{ij}(t, \Delta t)$ 可以视为一个关于年龄 t 的函数 $f_{ij}(t)$ 。若 t 时刻存在满足 $S(t)=i$ 的 n_i 个过程，以及 $t+\Delta t$ 时

刻该 n_i 个过程分别转移为四个状态的数量 $n_{i1}, n_{i2}, n_{i3}, n_{i4}$ ($\sum_{j=1}^4 n_{ij} = n_i$)。在局部中状态转移概率的无偏估计

即 $p_{ij}^*(t, \Delta t) = \frac{n_{ij}}{n_i}$ 。由于大数定律下 n_i 越大，该无偏估计结果的置信程度越高，状态转移概率模型更加适

合加权最小二乘法 (WLS) 对模型进行估计。

$$\hat{p}_{ij}(t) = f_{ij}(t) = \sum_{k=0} \beta_{ij,k} t^k \quad (\text{IV1})$$

由于状态转移数据能够提供每个 t 时刻处于 i 状态的 $n_{i,t}$ 个过程在 $t_n + \Delta t$ 转移时刻转移 j 状态的的过程个数 $n_{ij,t}$ 和未转移到 j 状态的过程个数 $n_{i,t} - n_{ij,t}$, 且每一个过程 S 是相互独立的, 该时刻的似然函数即可写作 $L(t; \beta_0, \beta_1, \dots, \beta_k) = \hat{p}_{ij}^{n_{ij,t}}(t) \cdot (1 - \hat{p}_{ij}(t))^{n_{i,t} - n_{ij,t}}$ 。在此基础上, 所有时刻 t 关于 $\hat{p}_{ij}(t)$ 的似然函数即可视作各时刻似然函数的乘积, 即式 (IV2)。

$$L(t, t + \Delta t, \dots; \beta_0, \beta_1, \dots, \beta_k) = \prod_{z=0} \hat{p}_{ij}^{n_{ij,t+z\Delta t}}(t + z\Delta t) \cdot (1 - \hat{p}_{ij}(t + z\Delta t))^{n_{i,t+z\Delta t} - n_{ij,t+z\Delta t}} \quad (\text{IV2})$$

$$\begin{aligned} \ln(L) &= \sum_{z=0} n_{i,t+z\Delta t} \left(\ln(\hat{p}_{ij}(t + z\Delta t)) + \frac{n_{ij,t+z\Delta t}}{n_{i,t+z\Delta t}} \ln(1 - \hat{p}_{ij}(t + z\Delta t)) \right) \\ &= \sum_{z=0} n_{i,t+z\Delta t} \left(\ln(\hat{p}_{ij}(t + z\Delta t)) + \hat{p}_{ij}^*(t + z\Delta t) \cdot \ln(1 - \hat{p}_{ij}(t + z\Delta t)) \right) \end{aligned} \quad (\text{IV3})$$

似然函数的对数形式即附录式 (IV3)。在参数估计中为使得似然函数最大, 每个时点的转移概率估计值 $\hat{p}_{ij}(t + z\Delta t)$ 应尽可能接近局部最优的无偏估计值 $\hat{p}_{ij}^*(t + z\Delta t)$, 其中 $n_{i,t+z\Delta t}$ 越大, 该局部的权重越大。因此, 本文选择采用每个时点的局部的无偏估计值 $\hat{p}_{ij}^*(t + z\Delta t)$ 作为作为模型拟合样本, 每个时点处于状态 i 的过程数量 $n_{i,t+z\Delta t}$ 作为权重, 通过加权最小二乘法 (WLS) 对模型参数进行估计, 进而得到标准状态转移概率模型 $p_{ij}(t, \Delta t)$ 和状态转移矩阵 $\mathbf{P}(t, \Delta t)$ 的最优估计式。本文模型的代码逻辑如表 A4 所示, 相关矩阵运算均通过 Python numpy 库实现。

$$\begin{aligned} \text{Loss} &= \sum_{z=0} n_{i,t+z\Delta t} (\hat{p}_{ij}(t + z\Delta t) - \hat{p}_{ij}^*(t + z\Delta t))^2 \\ &= \sum_{z=0} n_{i,t+z\Delta t} \left(\sum_{k=0} \beta_{ij,k} \cdot (t + \Delta t)^k - \hat{p}_{ij}^*(t + z\Delta t) \right)^2 \end{aligned} \quad (\text{IV4})$$

表 IV2 健康状态转移数据提取和处理过程

步骤 1: 提取每个参保人的状态变化时点的年龄 (以月为最小单位) 和对应状态, 直至数据截止时点。
步骤 2: 提取状态转移数据, 获得不同年龄和健康状态区间的状态转移数量。
步骤 3: 设定 t 的观测区间 $[t_{start}, t_{end}]$ (原文设置的为 50 岁-100 岁, 即拟合图中的 X 轴区间), 将状态转移间隔 Δt , 局部中状态转移概率的无偏估计即 $p_{ij}^*(t, \Delta t) = \frac{n_{ij}}{n_i}$ 。
步骤 4: 基于泰勒展开的思路, 每一个状态转移概率均可以用一个关于 t 的线性模型进行估计, 即 $\hat{p}_{ij}(t) = f_{ij}(t) = \sum_{k=0} \beta_{ij,k} t^k$, 权重为每个时点处于状态 i 的过程数量 $n_{i,t+z\Delta t}$ 。

步骤 5: 根据拟合后的模型, 分别提取 50-100 岁区间内 600 个月度时点的各个状态转移概率 p_{ij} , 组成 600 个状态转移概率矩阵 $\mathbf{P}(t, \Delta t)$, 根据

$$\mathbf{P}_{ij}(t, \Delta t_1 + \Delta t_2)_{(i,j \in \mathbf{I})} = \mathbf{P}_{ij}(t, \Delta t_1) \cdot \mathbf{P}_{ij}(t + \Delta t_1, \Delta t_2)_{(i,j \in \mathbf{I})}$$

进行矩阵连乘运算, 并从状态转移矩阵 $\mathbf{P}(t, \Delta t)$ 提取每一个状态转移概率模型 $p_{ij}(t, \Delta t)$ 。

步骤 6: 分别对 2017-2020 四个年度的状态数据重复以上步骤, 即可以比较各个年度的失能群体的健康状态转移概率 $p_{ij}^{year}(t, \Delta t)$ 。

附录 V 附表及附图

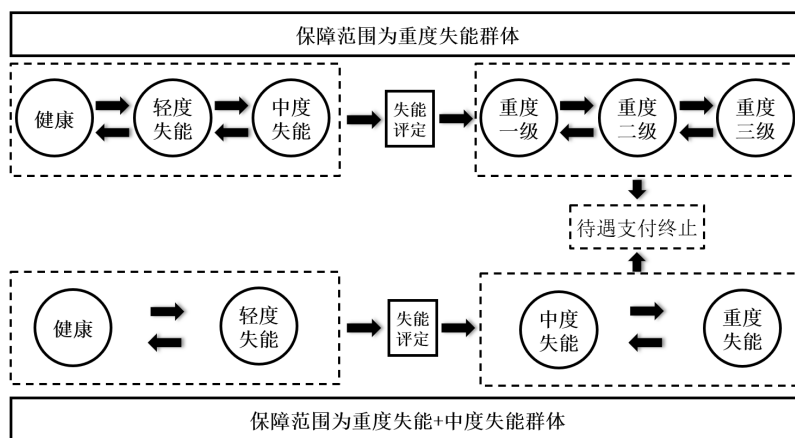


图 A1 长护险中的健康状态转移过程

表 A1 各年度失能建档人数及年龄结构的描述性统计

年度	评定等级	Count	Mean	Min	25%	50%	75%	Max	Std
2017	重度 1 级	4810	77.2	17	71	79	85	109	11.4
	重度 2 级	3898	80.5	24	75	82	87.75	109	10.6
	重度 3 级	72	79.6	41	73.75	81	89	108	13.3
2018	重度 1 级	8995	77.1	17	71	79	85	109	11.5
	重度 2 级	5000	80.0	20	74	82	88	105	11.1
	重度 3 级	71	77.7	40	72	79	86.5	108	13.2
2019	重度 1 级	9937	76.6	17	70	78	85	105	11.6
	重度 2 级	4697	79.3	20	73	81	87	106	11.6
	重度 3 级	80	77.5	40	72	80	86	96	12.4
2020	重度 1 级	9788	76.2	17	69	78	85	105	11.8
	重度 2 级	4293	78.7	20	73	81	87	106	11.8
	重度 3 级	140	77.6	40	71.75	80.5	86	99	12.5

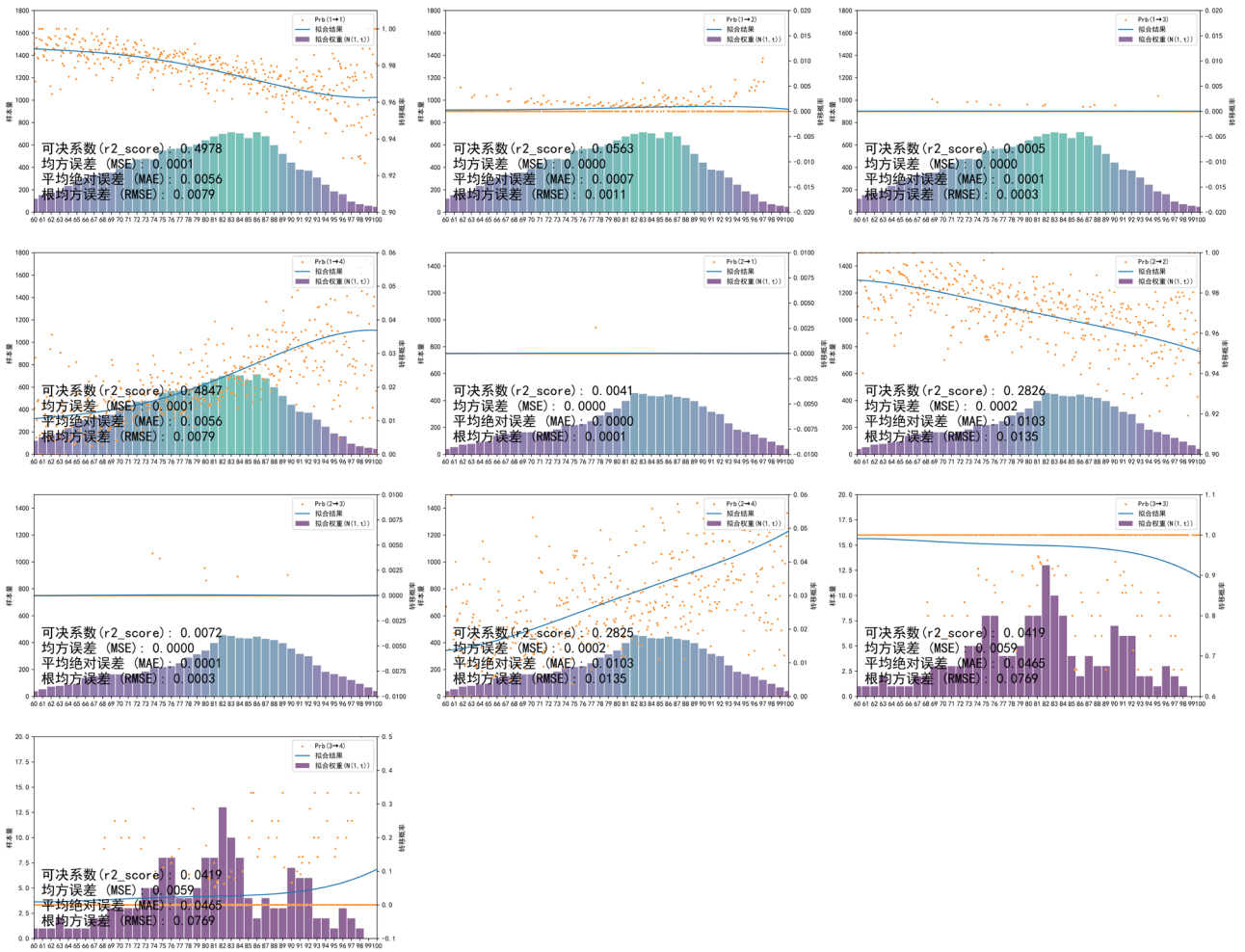


图 A2 总样本下各月度转移概率模型拟合效果和拟合参数

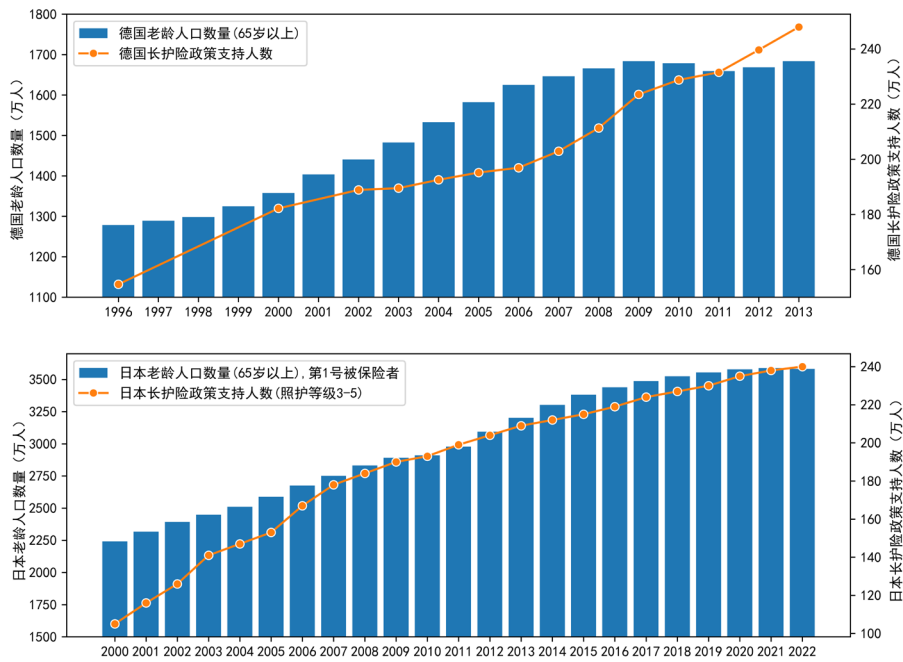


图 A3 德国和日本长护险制度建立初期的失能人数变化

表 A2 数学符号含义对照表

数学符号	含义	数学符号	含义
X	寿命随机变量	T	余寿随机变量
t	参保人年龄	Δt	t 的未来某时间段
i, j	多状态过程中转移前后健康状态	s	多状态过程中指参保人健康状态
l_t	活过 t 年龄的人数	$l_{t+\Delta t, u}$	Δt 时间后的失能人数
$p_{ij}(t, \Delta t)_{i, j \in \mathbf{I}}$	状态转移概率	$\mathbf{P}(t, \Delta t)$	状态转移矩阵
$\bar{q}_s(t, \Delta t)$	状态向量	\mathbf{I}	状态空间

注：该附录是期刊所发表论文的组成部分，同样视为作者公开发表的内容。如研究中使用该附录中的内容，请务必在研究成果上注明附录下载出处。