Теорія страхового та фінансового ризику

Ямненко Р.Є

КНУ імені Тараса Шевченка

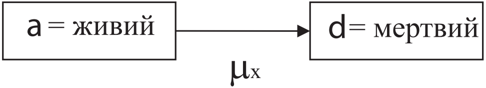
5 квітня 2017 р.

Зміст

1. Лекція 8. Модель Маркова з двома станами
   - Модель з двома станами (I) — припущення
   - Модель з двома станами (II) — ймовірності
   - Модель з двома станами (III) — статистика
   - Модель з двома станами (IV) — оцінка максимальної правдоподібності
   - Коментарі щодо застосування
   - Центральна експозиція ризику

Модель з двома станами (I) — припущення

1.1. Модель з двома станами проілюстровано на малюнку



У моделі два стани: "живий "мертвий" і можливий перехід тільки в одну сторону. Ймовірність того, що особа, яка дожила до даного віку x, помре у деякому наступному віці $x+t, \ (t \geq 0)$ визначається залежною від віку інтенсивністю переходу $\mu_{x+t}, \ (t \geq 0)$. При цьому нам будуть потрібні наступні припущення.

Припущення 1.

Ймовірність того, що особа даного віку x буде знаходитись у якомусь із двох станів у наступний час, залежить лише від віку x і від стану у якому вона знаходиться зараз.

Це марковське припущення.

Припущення 2.

$$ {}_{dt}q_{x+t} = \mu_{x+t}dt + o(dt), \;\; (t \geq 0). $$

1.2. Попередня історія особи, наприклад стан здоров'я, історія хвороби, місце роботи, виключаються із моделі. Якщо ми знаємо ці фактори, то ми можемо:

(a) розглядати кожну комбінацію факторів як окрему модель, іншими словами стратифікувати задачу;

(b) визначити модель так, щоб взяти до уваги всі фактори, іншими словами розглянути задачу як регресійну.

1.3. Для того,щоб робити висновки, ми обмежимося віком між x і $x+1$, і введемо таке припущення

Припущення 3.

$$\mu_{x+t} = \mu$$

для всіх $0 \leq t \leq 1$.

1.4. Важливо підкреслити, що ця модель з двома станами не співпадає із такою ж з лекції 5, бо ми маємо різні припущення. Одна модель сформульована у термінах випадкової величини T — тривалість життя, а друга в термінах інтенсивності переходу між станами. Легко ввести деякі слабкі умови при яких ці моделі будуть еквівалентними, але коли розглядаються більше ніж один приріст, то ці два формулювання ведуть у різних напрямках.

Модель з двома станами (II) — ймовірності

2.1. Так як ми визначаємо модель через інтенсивність переходу, то потрібно обчислити ймовірності переходу. Розглянемо ймовірність виживання ${}_{t+dt}p_x$ і умови на стан, який процес займає у віці $x+t$. На основі припущення марковості маємо

$$
{}_{t+dt}p_x =
$$

$$
= {}_t p_x \times P(\text{Живий у момент } x+t+dt \mid \text{Живий у момент } x+t) +
$$

$$
+ {}_t q_x \times P(\text{Живий у момент } x+t+dt \mid \text{Мертвий у момент } x+t) =
$$

$$
= {}_t p_x \times {}_{dt}p_{x+t} + {}_t q_x \times 0 = {}_t p_x \times (1 - \mu_{x+t}dt + o(dt)).
$$

Звідси

$$\frac{\partial}{\partial t}{}_tp_x = \lim_{dt\to 0+} \frac{{}_{t+dt}p_x - {}_tp_x}{dt}$$

$$= -{}_tp_x\mu_{x+t} + \lim_{dt\to 0+} \frac{o(dt)}{dt} = -{}_tp_x\mu_{x+t}. \qquad (*)$$

Отже,

$${}_tp_x = \exp\left\{-\int_0^t \mu_{x+s}ds\right\}$$

як і раніше.

2.2.

Важливо відмітити, що цей результат ми вивели прямо із зроблених для моделі з двома станами припущень, і що цей метод легко поширити на моделі з багатьма станами. У рамках теорії Маркова (*) є прикладом прямого рівняння Колмогорова.(див. лекцію 4).

Модель з двома станами (III) — статистика

3.1. Визначимо, яким чином проводяться спостереження. Припустимо, що проводиться спостереження за N особами протягом деякого скінченого періоду часу між віком x і віком $x+1$.

Будемо вважати N невипадковим і ми не будемо припускати ні те, що всі N осіб спостерігаються одночасно, ні те що кожна особа спостерігається протягом повного року життя.

Ми припускаємо, що тривалість життя всіх N осіб незалежні і однаково розподілені випадкові величини.

3.2.

Для простоти будемо розглядати цензурування типу I, але застосований підхід можна розповсюдити на більш реалістичні види цензурування. Для $i = 1, \ldots, N$ позначимо через

$$x + a_i$$

вік, з якого почалося спостереження за i-ю особою, а через

$$x + b_i$$

позначимо вік, коли спостереження за i-ю особою припиняються, якщо особа дожила до цього віку. $x + b_i$ може дорівнювати або $x + 1$, або віку i-ї особи, якщо спостереження припинилися раніше.

3.3.

Визначимо випадкову величину

$$\mathbf{D}_i = \begin{cases} 1, & \text{якщо спостерігається смерть } i\text{-тої особи;} \\ 0, & \text{якщо не спостерігається смерть } i\text{-тої особи.} \end{cases}$$

3.4. $\mathbf{D}_i$ є прикладом індикаторної випадкової величини, вона відмічає смерть i-тої особи. Визначимо випадкову величину $\mathbf{T}_i$ так

$x+\mathbf{T}_i =$ вік, коли спостереження за i-тою особою припиняються.

Зауважимо, що випадкові величини $\mathbf{D}_i$ і $\mathbf{T}_i$ не є незалежними тому, що:

$$\mathbf{D}_i = 0 \;\Leftrightarrow\; \mathbf{T}_i = b_i, \quad \mathbf{D}_i = 1 \;\Leftrightarrow\; a_i < \mathbf{T}_i < b_i.$$

3.5. Часто буває зручно мати справу з часом, протягом якого проводились спостереження за i-тою особою, тому визначимо випадкову величину

$$\mathbf{V}_i = \mathbf{T}_i - a_i.$$

$\mathbf{V}_i$ називають *часом чекання (waiting time)*.

Ця випадкова величина має розподіл змішаного типу із стрибком функції розподілу у точці $b_i - a_i$.

3.6.

Пара

$$(\mathbf{D}_i, \mathbf{V}_i)$$

утворює статистику у тому розумінні, що результатом нашого спостереження є вибірка (d_i, v_i) із розподілу $(\mathbf{D}_i, \mathbf{V}_i)$.

Нехай $f_i(d_i, v_i)$ – це сумісний розподіл $(\mathbf{D}_i, \mathbf{V}_i)$. Його легко записати розглянувши два випадки $\mathbf{D}_i = 0$ і $\mathbf{D}_i = 1$.

$$f_i(d_i, v_i) = \begin{cases} {}_{b_i - a_i}p_{x+a_i}, & (d_i = 0) \\ {}_{v_i}p_{x+a_i}\, \mu_{x+a_i+v_i}, & (d_i = 1) \end{cases}$$

$$= \begin{cases} \exp\left\{-\int_0^{b_i - a_i} \mu_{x+a_i+t}\, dt\right\}, & (d_i = 0) \\ \exp\left\{-\int_0^{v_i} \mu_{x+a_i+t}\, dt\right\} \mu_{x+a_i+v_i}, & (d_i = 1) \end{cases} =$$

$$= \exp\left\{-\int_0^{v_i} \mu_{x+a_i+t}\, dt\right\} \mu_{x+a_i+v_i}^{d_i}.$$

Тепер використаємо припущення, що $\mu_{x+t} = \mu$ при $0 \le t < 1$, і одержимо

$$f_i(d_i, v_i) = e^{-\mu v_i} \mu^{d_i}.$$

3.7.

Оскільки всі $(\mathbf{D}_i, \mathbf{V}_i)$ незалежні, то сумісний розподіл має вигляд

$$\prod_{i=1}^{N} e^{-\mu v_i} \mu^{d_i}$$

$$= e^{-\mu(v_1+\ldots+v_N)} \mu^{d_1+\ldots+d_N} = e^{-\mu v} \mu^{d},$$

де $d = \sum_{i=1}^{N} d_i$ і $v = \sum_{i=1}^{N} v_i$.

Якщо визначити випадкові величини $\mathbf{D}$ і $\mathbf{V}$ відповідно, як загальну кількість смертей і загальний час чекання, тоді сумісний розподіл всіх $(\mathbf{D}_i, \mathbf{V}_i)$ легко виражається в термінах $\mathbf{D}$ і $\mathbf{V}$.

Модель з двома станами (IV) — оцінка максимальної правдоподібності

4.1. Функція правдоподібності має вигляд

$$L(\mu; d, v) = e^{-\mu v}\mu^{d},$$

звідки одержимо оцінку максимальної правдоподібності для μ:

$$\hat{\mu} = d/v.$$

4.2. Значення оцінки $\hat{\mu}$ є функцією вибіркових значень d і v, і це значення може розглядатись, як вибірка із розподілу відповідної оцінки

$$\tilde{\mu} = \mathbf{D}/\mathbf{V}.$$

4.3.

Ми будемо використовувати для позначення випадкових величин великі жирні літери, а для позначення вибірки – малі літери.

4.4. У застосуваннях важливо уміти оцінювати моменти оцінки $\tilde{\mu}$, наприклад для порівняння результатів дослідження із стандартними таблицями. Принаймні нам необхідно оцінити $E[\tilde{\mu}]$ і $\mathrm{Var}[\tilde{\mu}]$.

Мають місце такі точні результати

$$E[\mathbf{D}_i - \mu \mathbf{V}_i] = 0, \tag{1}$$

$$\mathrm{Var}[\mathbf{D}_i - \mu \mathbf{V}_i] = E[\mathbf{D}_i]. \tag{2}$$

Зауважимо, що перше співвідношення можна переписати у вигляді

$$E[\mathbf{D}_i] = \mu E[\mathbf{V}_i].$$

Доведемо співвідношення (1), (2).

Маємо

$$E[\mathrm{D}_i] = P(\mathrm{D}_i = 1) = 1 - P(\mathrm{D}_i = 0) = 1 - {}_{b_i - a_i}p_{x+a_i}$$

$$= 1 - \exp\left\{ -\int_0^{b_i - a_i} \mu \, dt \right\} =$$

$$= 1 - \exp\{-\mu(b_i - a_i)\}.$$

З другого боку

$$E[\mathbf{V}_i] = \int_0^{b_i-a_i} t \cdot {}_t p_{x+a_i} \cdot \mu_{x+a_i+t}\, dt + (b_i - a_i) \cdot {}_{b_i-a_i} p_{x+a_i}$$

$$= \int_0^{b_i-a_i} t e^{-\mu t} \mu\, dt + (b_i - a_i) e^{-\mu(b_i-a_i)} =$$

$$= \left(-t e^{-\mu t}\right)\Big|_0^{b_i-a_i} + \int_0^{b_i-a_i} e^{-\mu t}\, dt + (b_i - a_i) e^{-\mu(b_i-a_i)}$$

$$= \frac{1}{\mu}\left(1 - e^{-\mu(b_i-a_i)}\right) = \frac{1}{\mu} E[\mathrm{D}_i].$$

Використаємо розподіл $f_i(d_i, v_i) = e^{-\mu v_i}\mu^{d_i}$. Маємо

$$\mathrm{Var}[\mathsf{D}_i - \mu\mathbf{V}_i] = E[(\mathsf{D}_i - \mu\mathbf{V}_i)^2] - (E[\mathsf{D}_i - \mu\mathbf{V}_i])^2$$

$$= \int_0^{b_i - a_i} (1 - \mu v)^2 \mu e^{-\mu v}\, dv +$$

$$+(0 - \mu(b_i - a_i))^2 P(\mathsf{D}_i = 0) = \left(1 - e^{-\mu(b_i - a_i)}\right) = E[\mathsf{D}_i].$$

Щоб знайти асимптотичний розподіл $\tilde{\mu}$, розглянемо

$$\frac{1}{N}(\mathsf{D} - \mu\mathbf{V}) = \frac{1}{N}\sum_{i=1}^{N}(\mathsf{D}_i - \mu\mathbf{V}_i).$$

Далі зауважимо, що (не строго)

$$\lim_{N\to\infty}(\tilde{\mu} - \mu) = \lim_{N\to\infty}\frac{N}{\mathbf{V}}\left(\frac{\mathsf{D}}{N} - \frac{\mu\mathbf{V}}{N}\right).$$

За законом великих чисел $\mathbf{V}/N \to E[\mathbf{V}_i]$, і за центральною граничною теоремою

$$\frac{1}{N}(\mathsf{D} - \mu\mathbf{V}) \sim \mathrm{Normal}\left(0, \frac{E[\mathsf{D}]}{N^2}\right),$$

тому асимптотично

$$\tilde{\mu} \sim \mathrm{Normal}\left(\mu, \frac{\mu}{E[\mathbf{V}]}\right).$$

Коментарі щодо застосування

5.1. Маючи оцінену кусково-сталу інтенсивність (сталу протягом одного року віку), ми можемо використати цей результат для оцінки функції μ_x, як гладкої функції віку.

Процес згладжування називається *градуюванням (graduation)*.

Для цього зазвичай припускають, що $\hat{\mu}$ оцінює

$$\mu_{x+1/2}.$$

5.2.

Використавши чисельні методи, ми можемо обчислити всі потрібні нам ймовірності із

$$ {}_t p_x = \exp\left\{-\int_0^t \mu_{x+s}\, ds\right\}. $$

Центральна експозиція ризику

6.1. У актуарній термінології спостережений час чекання у віці x, який ми позначили через v, часто називають *центральною експозицією ризику* і позначають

$$E_x^c.$$