Ανάλυση Επιβίωσης

Επικ. Καθ. Σ. Ζημερας
Τμήμα Στατιστικής και Αναλογιστικών – Χρηματοοικονομικών Μαθηματικών
Πανεπιστήμιο Αιγαίου
Σάμος
2020

ΥΠΟ ΣΥΝΘΗΚΗ ΠΙΘΑΝΟΤΗΤΕΣ ΘΑΝΑΤΟΥ

- Η υπο-συνθηκη πιθανότητα θανάτου στο διάστημα ηλικιών (x,x+t) δοθέντα ότι το άτομο βρίσκεται στην ζωή κατά την ηλικία x, παίζει σημαντικό ρόλο.
- Στα ασφαλιστικά μαθηματικά συμβολίζεται με ${}_tq_x$ όπου ο δείκτης t δηλώνει το μήκος της περιόδου για την οποία υπολογίζεται αυτή η πιθανότητα.
- Έχουμε

$$ {}_tq_x = p\left(x < X < x+t \mid X > x\right) = \frac{\int_x^{x+t} f_x(u)\,du}{\int_x^{+\infty} f_x(u)\,du} = \frac{\int_x^{x+t} h_x(u) S_x(u)\,du}{\int_x^{+\infty} h_x(u) S_x(u)\,du} = \frac{S_x(x) - S_x(x+t)}{S_x(x)} $$

ΥΠΟ ΣΥΝΘΗΚΗ ΠΙΘΑΝΟΤΗΤΕΣ ΘΑΝΑΤΟΥ

Η διαφορά S(x)-S(x+t) είναι η αναλογία θανάτου μεταξύ των ηλικιών x και x+t.

Ο κεντρικός ρυθμός που σημειώνεται με t^{mx} είναι

$$t^{m_x} = \frac{\int_x^{x+t} h_x(u) S_x(u)\,du}{\int_x^{x+t} S_x(u)\,du} = \frac{S_x(x) - S_x(x+t)}{\int_x^{x+t} S_x(u)\,du}$$

Η υπο-συνθηκη πιθανότητα επιβίωσης στο διάστημα (x,x+t) δοθέντος ότι το άτομο βρίσκεται στην ζωή κατά την ηλικία x είναι

$${}_t p_x = 1 - {}_t q_x = \frac{S_x(x+t)}{S_x(x)}$$

ΥΠΟ ΣΥΝΘΗΚΗ ΠΙΘΑΝΟΤΗΤΕΣ ΘΑΝΑΤΟΥ

- Για σύνολο σημείων $x_0 = 0 < x_1 < x_2 < \ldots < x_k$ με $x_{i+1} - x_i = h_i$ έχουμε

$$S_x\left(x_k\right) = e^{\int_0^{x_k} h(u)du} = e^{-\sum_{i=0}^{k-1}\int_{x_i}^{x_{i+1}} h_x(u)du} = \prod_{i=0}^{k-1} e^{-\int_{x_i}^{x_{i+1}} h(u)du} = \prod_{i=0}^{k-1} {}_{h_i}p_{x_i} = \prod_{i=0}^{k-1}\left(1 - {}_{h_i}q_{x_i}\right)$$

Η έκφραση είναι χρήσιμη για την κατασκευή πινάκων επιβίωσης.

ΠΕΡΙΚΟΜΜΕΝΕΣ ΚΑΤΑΝΟΜΕΣ

- Δεδομένα θνησιμότητας συλλέγονται συχνά μέσα σε συγκεκριμένο χρονικό διάστημα. Επομένως η χρήση περικομμμένων κατανομών δίνεται σε αυτές τις περιπτώσεις.
- Έστω ότι μπορούμε να παρατηρήσουμε την τ.μ. T σε πεπερασμένο διάστημα α<t<b.

$$F_T(t \mid a < t < b) = \begin{cases} 0, t \le a \\ \dfrac{F_T(t) - F_T(a)}{F_T(b) - F_T(a)}, a < t \le b \\ 1, t > b \end{cases}$$

ΠΕΡΙΚΟΜΜΕΝΕΣ ΚΑΤΑΝΟΜΕΣ

- Αν χρησιμοποιήσουμε την $S_T(t)$ θα έχουμε

$$S_T(t \mid a < t < b) = \begin{cases} 0, t \le a \\ \dfrac{S_T(t) - S_T(b)}{S_T(a) - S_T(b)}, a < t \le b \\ 1, t > b \end{cases}$$

- Αν η αρχική κατανομή της Τ έχει σ.π.π $f_T(t)$ τότε η περικομμένη κατανομή έχει σ.π.π που διωεται από την σχέση

$$f_T(t \mid a < t < b) = \begin{cases} \dfrac{f_T(t)}{F_T(b) - F_T(a)} = \dfrac{f_T(t)}{S_T(a) - S_T(b)}, a < t \le b \\ 0, \forall t \in (-\infty, a] \cup (b, +\infty) \end{cases}$$

ΠΕΡΙΚΟΜΜΕΝΕΣ ΚΑΤΑΝΟΜΕΣ

- Η συνάρτηση κινδύνου είναι

$$h_T(t \mid a<T<b) = \frac{f_T(t \mid a<T<b)}{S_T(t \mid a<T<b)} = \frac{f_T(t)}{S_T(t) - S_T(b)}$$

- Η τελευταία σχέση μπορεί να γραφτεί

$$h_T(t \mid a<T<b) = \frac{f_T(t)}{S_T(t) - S_T(b)} \frac{S_T(t)}{S_T(t) - S_T(b)} = h_T(t) \frac{S_T(t)}{S_T(t) - S_T(b)}$$

- Ο λόγος των συναρτήσεων κινδύνου για τις περικομμένες και μη περικομμένες κατανομές εξαρτάται **μόνο από το b** που αποκαλείται και άνω σημείο περικοπής και όχι από το α γνωστό σαν κάτω σημείο περικοπής.

ΠΕΡΙΚΟΜΜΕΝΕΣ ΚΑΤΑΝΟΜΕΣ

- Αν η κατανομή είναι περικομμένη από κάτω παίρνουμε $b=+\infty$ και ισχύει $S_T(b)=0$, οπότε

$$h_T\left(t \mid T > a\right) = h_T\left(t\right)$$

- Η περικοπή από αριστερά δεν μεταβάλλει την συνάρτηση κινδύνου ενώ η προς τα δεξιά την αυξάνει.

ΠΕΡΙΚΟΜΜΕΝΕΣ ΚΑΤΑΝΟΜΕΣ

- Μέση τιμή της τ.μ. Τ

$$E(T)=\int_{-\infty}^{+\infty} t f_T(t)\,dt=\int_{-\infty}^{+\infty} t h_T(t) S_T(t)\,dt$$

- Διασπορά της τ.μ. Τ

$$V(T)=E\left\{\left[T-E(T)\right]^2\right\}=E\left(T^2\right)-\left\{E(T)\right\}^2=\int_{-\infty}^{+\infty} t^2 f_T(t)\,dt-\left\{E(T)\right\}^2$$

ΠΕΡΙΚΟΜΜΕΝΕΣ ΚΑΤΑΝΟΜΕΣ

- Αν X τ.μ. που παριστάνει την ηλικία.
- Εισάγουμε μια νέα μεταβλητή T που παριστάνει τον χρόνο που έχει διαρρεύσει μετά την ηλικία X. Η υπο-συνθήκη κατανομή του T δοθέντος του x είναι η κατανομή του χρόνου της μελλοντικής ζωής
- Η σ.π.π είναι

$$f_T(t) = \frac{f_x(x+t)}{S_x(x)}, t > 0$$

- Η συνάρτηση επιβίωσης είναι

$$S_T(t) = \frac{S_x(x+t)}{S_x(x)}, t > 0$$

ΠΕΡΙΚΟΜΜΕΝΕΣ ΚΑΤΑΝΟΜΕΣ

- Η συνάρτηση κινδύνου είναι

$$h_T(t) = h_x(x+t), t > 0$$

- Η αναμενόμενη τιμή του Τ καλείται η **αναμενόμενη τιμή του χρόνου μελλοντικής ζωής** ενός ατόμου ηλικίας x και συμβολίζεται με e^0_x με

$$e_x^0 = E(T) = E(X - x \mid X > x) = \int_0^{+\infty} t f_T(t)\,dt = \int_0^{+\infty} t h_T(t) S_T(t)\,dt$$

Αλλά $h_T(t) = h_x(x+t)$ και $f_T(t) = \dfrac{f_x(x+t)}{S_x(x)}$

Επομένως

$$e_x^0 = \frac{1}{S_x(x)} \int_0^{+\infty} t f_X(x+t)\,dt = \frac{1}{S_x(x)} \int_0^{+\infty} t h_X(x+t) S_X(x+t)\,dt$$

ΠΕΡΙΚΟΜΜΕΝΕΣ ΚΑΤΑΝΟΜΕΣ

- Η μέση ηλικία θανάτου είναι

$$x + e_x^0$$

- Η διακύμανση του T είναι

$$V(T) = V(X \mid X > x) = \int_0^{+\infty} t^2 f_T(t)\,dt - \left(e_x^0\right)^2 = \frac{1}{S_x(x)} \int_0^{+\infty} t^2 f_X(x+t)\,dt - \left(e_x^0\right)^2$$