Ανάλυση Επιβίωσης

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

ΜΕΘΟΔΟΣ NELSON

- Θα ορίσουμε αρχικά την εμπειρική συνάρτηση επιβίωσης
- Έστω $t_{(1)} < t_{(2)} < \ldots .< t_{(n)}$ η διακεκριμένοι και διατεταγμένοι χρόνοι θανάτου. Άρα η θεωρητική αθροιστική κατανομή είναι

$$p\left(T \leq t\right) = F\left(t\right)$$

και η αντίστοιχη συνάρτηση επιβίωσης

$$p\left(T > t\right) = 1 - F\left(t\right) = S\left(t\right)$$

- Η εμπειρική κατανομή θανάτου είναι

$$F_n^0\left(t\right) = \begin{cases} 0, t < t_{(1)} \\ \dfrac{i}{n} \quad , t_{(i)} < t \leq t_{(i+1)} \\ 1, t > t_{(n)} \end{cases}$$

ΜΕΘΟΔΟΣ NELSON

Η εμπειρική συνάρτηση επιβίωσης είναι

$$S_n^0(t) = 1 - F_n^0(t)$$

και

$$S_n^0(t) = \begin{cases} 1, t < t_{(1)} \\ \dfrac{n-i}{n} \quad , t_{(i)} < t \le t_{(i+1)} \\ 0, t > t_{(n)} \end{cases}$$

Οι συναρτήσεις $F^0{}_n$ (t) και $S^0{}_n$ (t) είναι συνεχείς από τα δεξιά και αποτελούν εκτιμήσεις των p(T<t) και p(T>t) αντίστοιχα.

ΜΕΘΟΔΟΣ NELSON

Άρα $$S_n^0(t)=\frac{n-i}{n}=\frac{n-1}{n}\frac{n-2}{n-1}\cdots\frac{n-i}{n-i+1}=\left(1-\frac{1}{n}\right)\left(1-\frac{1}{n-1}\right)\cdots\left(1-\frac{1}{n-i+1}\right)$$

όπου $T\in[t_{(i)}, t_{(i+1)}]$

Είναι γνωστό ότι $H(t)=-\ln S(t)$ έτσι για την εμπειρική $H^0_n(t)$ έχουμε

$$H_n^0(t)=-\ln S_n^0(t)=\sum_{j=1}^{i}\ln\left(1-\frac{1}{n-j+1}\right)$$

όπου $t\in[t_{(i)}, t_{(i+1)}]$, i=1,2,…,n

Για μεγάλο n-j+1 μπορούμε να χρησιμοποιήσουμε την σχέση

$$\ln\left(1-\frac{1}{n-j+1}\right)\simeq\frac{1}{n-j+1}$$

ΜΕΘΟΔΟΣ NELSON

- Έτσι έχουμε

$$H_n^0(t) = -\ln S_n^0(t) \simeq \sum_{j=1}^{i} \frac{1}{n-j+1} = \frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{n-j+1}$$

- Επομένως η εμπειρική αθροιστική συνάρτηση κινδύνου μπορεί να εκφραστεί προσεγγιστικά ως το άθροισμα των τάξεων των χρόνων θανάτου.
- Αν έχουμε αθροιστική συνάρτηση κινδύνου μέχρι το χρονικό σημείο $t_{(i)}$ τότε

$$0 < H^0\left(t_{(1)}\right) < H^0\left(t_{(2)}\right) < \ldots < H^0\left(t_{(i)}\right)$$

- Ο Nelson (1972) έδηξε ότι

$$E\left[H^0\left(T_i\right)\right] = \frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{n-i+1} = \overline{H^0\left(t_i\right)}$$

ΜΕΘΟΔΟΣ NELSON

Βήματα

- Γράφουμε τους χρόνους θανάτου $t_{(i)}$ σε φθίνουσα σειρά
- Γράφουμε τις τάξεις των παρατηρήσεων με αντίστροφη σειρά
- Υπολογίζουμε τα $(n\text{-}i+1)^{-1}$ και την κατά προσέγγιση εμπειρική $H^0{}_n(t)$, χρησιμοποιώντας την

$$H_n^0(T_i) = \sum_{j=1}^{i} \frac{1}{n-j+1}$$

  για κάθε σημείο t_i
- Κάνουμε την γραφική παράσταση μιας κατάλληλης συνάρτησης της $H^0{}_n(t)$ ως προς μια κατάλληλη συνάρτηση του t για να βρούμε μια ευθεία γραμμή

ΜΕΘΟΔΟΣ NELSON

- Παράδειγμα

Χρόνοι θανάτου 22 ποντικιών που πέθαναν από λέμφωμα

i	t_i	$S_n^0(t_i)$	$H_n^0(t_i) = -\ln S_n^0(t_i)$	$(n-i+1)^{-1}$	$\simeq H_n^0(t_i)$
1	159	0.954	0.046	0.045	0.045
2	189	0.909	0.095	0.047	0.093
3	191	0.863	0.146	0.05	0.143
⋮	⋮	⋮	⋮	⋮	⋮
20	414	0.09	2.39	0.33	2.18
21	428	0.045	3.09	0.5	2.68
22	432	-	-	1	3.68

$$S_n^0(t) = \begin{cases} 1, t < t_{(1)} \\ \dfrac{n-i}{n} \quad , t_{(i)} < t \leq t_{(i+1)} \\ 0, t > t_{(n)} \end{cases}$$

$$S_1^0(t) = \frac{n-i}{n} = \frac{22-1}{22} = 0.954$$

$$H_n^0(t) = -\ln S_n^0(t) \simeq \sum_{j=1}^{i} \frac{1}{n-j+1} = \frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{n-j+1}$$

$$H_1^0(t) = \frac{1}{n} = \frac{1}{22} = 0.045$$

ΜΕΘΟΔΟΣ NELSON

- Ο προηγούμενος πίνακας δίνει τους υπολογισμούς της συνάρτησης κινδύνου. Δίνονται 2 τρόποι
- Στην στήλη 3 δίνεται η εμπειρική συνάρτηση επιβίωσης με βάση τον τύπο

$$S_n^0(t) = \frac{n-i}{n} = \frac{n-1}{n}\frac{n-2}{n-1}\cdots\frac{n-i}{n-i+1}$$

- Στην στήλη 4 η εμπειρική συνάρτηση κινδύνου υπολογίζεται με βάση τον τύπο

$$H_n^0(t) = -\ln S_n^0(t)$$

- Η τελευταία στήλη υπολογίζεται από τον τύπο

$$H_n^0(t) \simeq \sum_{j=1}^{i}\frac{1}{n-j+1}$$

ΜΕΘΟΔΟΣ NELSON

- Αν θέλουμε να χρησιμοποιήσουμε την κατανομή Gompertz, η αθροιστική συνάρτηση κινδύνου της κατανομής Gompertz δίνεται από

$$H(t) = \frac{k}{a}\left(e^{at} - 1\right)$$

- Η μέθοδος Nelson επιβάλλει να υπολογίσουμε την γραφική παράσταση μιας συνάρτησης της

$$H_n^0(t) \simeq \sum_{j=1}^{i} \frac{1}{n - j + 1}$$

για να βρούμε μια ευθεία γραμμή

ΜΕΘΟΔΟΣ NELSON

- Εναλλακτική εφαρμογή της κατανομής Gompertz
- Αν x=0 η αναμενόμενη ζωή και διασπορά δίνονται από τις σχέσεις

$$e_0^0 = \frac{1}{S_x(0)} \int_0^{+\infty} t f_X(t)\,dt = \frac{1}{S_x(0)} \int_0^{+\infty} t h_X(t) S_X(t)\,dt$$

- Η μέση ζωή κατά την στιγμή του θανάτου είναι $x + e_x^0 = 0 + e_0^0$
- Η διασπορά δίνεται από

$$V(T) = V(X \mid X > 0) = V(X) = \int_0^{+\infty} t^2 f_T(t)\,dt - \left(e_0^0\right)^2 = \frac{1}{S_x(0)} \int_0^{+\infty} t^2 f_X(t)\,dt - \left(e_0^0\right)^2$$

- Επειδή

$$S_x(0) = 1 \Rightarrow V(T) = \int_0^{+\infty} t^2 f_x(t)\,dt - \left(e_0^0\right)^2$$

με

$$e_0^0 = \int_0^{+\infty} t h_X(t) S_X(t)\,dt$$