Homework 9

Yuki Joyama

2024-04-23

1

$$h(x) = \frac{2x}{(1+x^2)}$$

Let $H(x)$ be cumulative hazard function, $S(x)$ be Survival function. and $f(x)$ be density function.

$$H(x) = \int_0^x h(t)\,dt$$
$$= \int_0^x \frac{2}{1+t^2}\,dt = \left[\log(1+t^2)\right]_0^x$$
$$= \log(1+x^2)$$

$$S(x) = \exp\{-H(x)\} = \exp(-\log(1+x^2)) = \underline{\frac{1}{1+x^2}}$$

$$h(x) = \frac{f(x)}{S(x)} \Rightarrow f(x) = h(x) \times S(x) = \frac{2x}{(1+x^2)} \times \frac{1}{(1+x^2)}$$
$$= \underline{\frac{2x}{(1+x^2)^2}}$$

2

t_i	n_i	d_i	c_i	$\hat{\lambda}_i = \frac{d_i}{n_i}$	$\hat{S}(t)$
1	10	1	0	1/10	$1 \times (1 - 1/10) = 9/10$
2	9	2	0	2/9	$9/10 \times (1 - 2/9) = 7/10$
4	7	0	1	0/7	$7/10 \times (1 - 0/7) = 7/10$
5	6	0	1	0/6	$7/10 \times (1 - 0/6) = 7/10$
6	5	1	0	1/5	$7/10 \times (1 - 1/5) = 14/25$
7	4	0	1	0/4	$14/25 \times (1 - 0/4) = 14/25$
8	3	0	1	0/3	$14/25 \times (1 - 0/3) = 14/25$
9	2	0	1	0/2	$14/25 \times (1 - 0/2) = 14/25$
10	1	0	1	0/1	$14/25 \times (1 - 0/1) = 14/25$

t_i: time to event
n_i: number of observations at t_i
d_i: number of events at t_i
c_i: number of censored observations at t_i
$\hat{\lambda}_i$: conditional probability of event at t_i given that the individual is still alive at t_i
$\hat{S}(t)$: Kaplan-Meier estimator of survival function

(a) Kaplan-Meier estimate of the survival function

$$\hat{S}(t) = \prod_{i=1}^{k} (1 - \hat{\lambda}_i) = \prod_{i=1}^{k} (1 - d_i/n_i), \quad t_k \le t < t_{k+1}$$

The values for each $t_k \le t < t_{k+1}$ $(k = 1, 2, \cdots, 10)$ are stated in the above table.

(b) Nelson-Aalen estimate of the cumulative hazard function

$$\tilde{H}(t) = \begin{cases} 0 & 0 \le t < t_1 \\ \sum_{t_i \le t} d_i/n_i & t \ge t_1 \end{cases} \Rightarrow$$

t_i	d_i/n_i	$\tilde{H}(t)$
1	1/10	1/10 = 0.1
2	2/9	1/10 + 2/9 = 0.32
4	0/7	0.32
5	0/6	0.32
6	1/5	0.32 + 1/5 = 0.52
7	0/4	0.52
8	0/3	0.52
9	0/2	0.52
10	0/1	0.52

(c) Fleming-Harrington estimate of the survival function

t_i	$\exp(-\tilde{H}(t))$
1	0.905
2	0.726
4	0.726
5	0.726
6	0.595
7	0.595
8	0.595
9	0.595
10	0.595

3

`tongue` data contains the following columns:
- `type` Tumor DNA profile (1 = Aneuploid Tumor, 2 = Diploid Tumor)
- `time` Time to death or on-study time, weeks
- `delta` Death indicator (0 = alive, 1 = dead)

Here we consider individuals with `delta` = 0 as right censored ones.

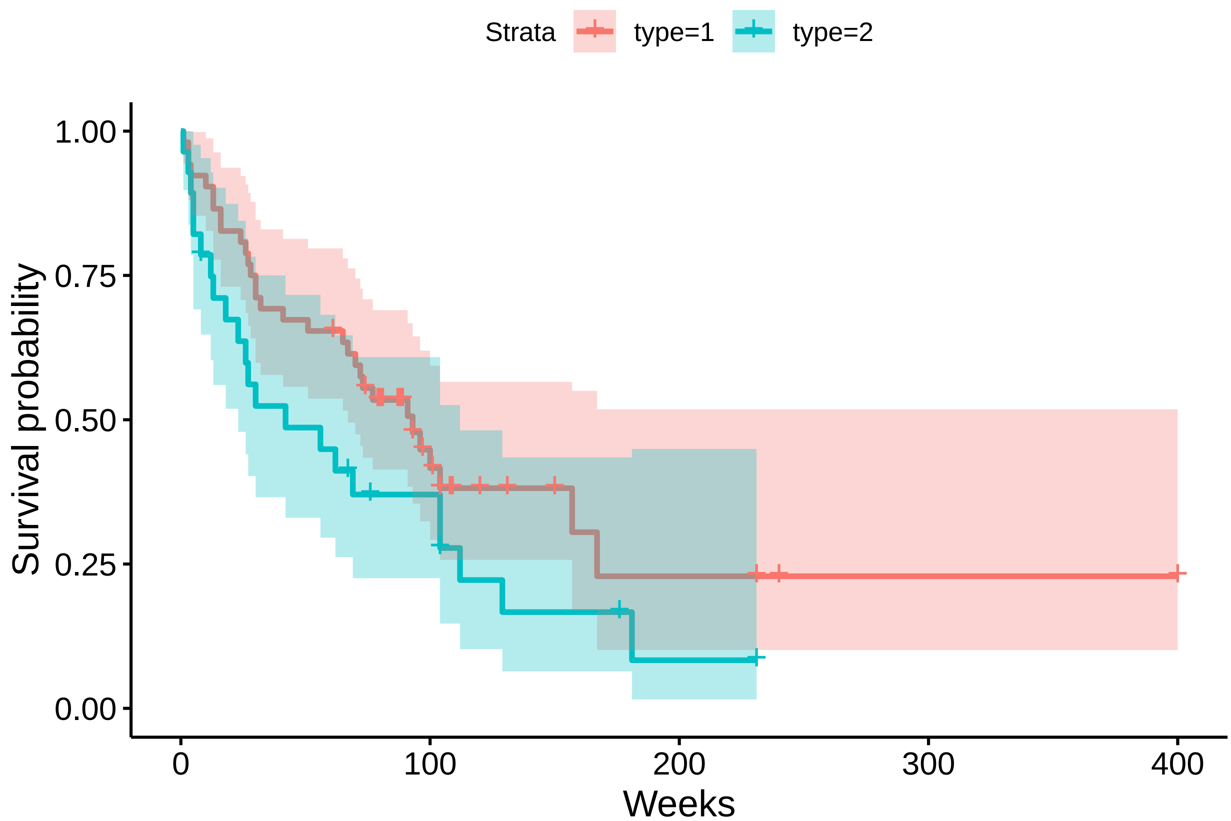
```
# data import
data("tongue")

str(tongue)
```

```
## 'data.frame':    80 obs. of  3 variables:
##  $ type : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ time : int  1 3 3 4 10 13 13 16 16 24 ...
##  $ delta: int  1 1 1 1 1 1 1 1 1 1 ...
```

```
# check censored
Surv(tongue$time, tongue$delta, type = "right")
```

```
##  [1]   1    3    3    4   10   13   13   16   16   24   26   27   28   30   30
## [16]  32   41   51   65   67   70   72   73   77   91   93   96  100  104  157
## [31] 167   61+  74+  79+  80+  81+  87+  87+  88+  89+  93+  97+ 101+ 104+ 108+
## [46] 109+ 120+ 131+ 150+ 231+ 240+ 400+   1    3    4    5    5    8   12   13
## [61]  18   23   26   27   30   42   56   62   69  104  104  112  129  181    8+
## [76]  67+  76+ 104+ 176+ 231+
```

```
# plot Kaplan-Meier curve of survival function
ggsurvplot(survfit(Surv(time, delta) ~ type, data = tongue, conf.type = "log"), conf.int = TRUE, xlab =
```



The Kaplan-Meier curve of survival function and its pointwise 95% CI using the log transformation is shown above.

```
# estimated 1-year (52 weeks) survival rate and 95% CI
KM <- survfit(Surv(time, delta) ~ type, data = tongue, conf.type = "log")

summary(KM, times = 52)
```

```
## Call: survfit(formula = Surv(time, delta) ~ type, data = tongue, conf.type = "log")
##
##                 type=1
##         time       n.risk       n.event      survival      std.err lower 95% CI
##       52.000       34.000        18.000         0.654        0.066        0.537
## upper 95% CI
##        0.797
##
##                 type=2
##         time       n.risk       n.event      survival      std.err lower 95% CI
##      52.0000      13.0000       14.0000        0.4864       0.0961       0.3302
## upper 95% CI
##       0.7164
```

Given the output, individuals with Aneuploid Tumor (`type` = 1) have an estimated 1-year (52 weeks) survival rate of 0.654 (95% CI: 0.537 - 0.797) and individuals with Diploid Tumor (`type` = 2) have that of 0.4864 (95% CI: 0.3302 - 0.7164).