HW9 BMII

Shihui Zhu

1. $h(x) = \frac{2x}{(1+x^2)} = -\frac{\partial \log(S(x))}{\partial x} = \frac{f(x)}{S(x)}$

$-\int_0^x h(t)dt = -\int_0^x \frac{2t}{(1+t^2)}dt = -\int_0^x \frac{1}{u}du = -\log u\Big|_0^x = -\log(1+t^2)\Big|_0^x = -\log(1+x^2)$

↳ $u = 1+t^2 \Rightarrow \frac{du}{dt} = 2t \Rightarrow du = 2t \cdot dt$

$S(x) = \exp(-\log(1+x^2))$

$= \frac{1}{1+x^2}$

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

⇒ Therefore, the survival function is $S(x) = \frac{1}{1+x^2}$, the density function is $f(x) = \frac{2x}{(1+x^2)^2}$

2.

					KM	Nelso-Aalen	Fleming-Harrington
t_i	n_i	d_i	c_i	$\hat{\lambda}_i$	$\hat{S}(t)$	$\hat{H}(t)$	$\exp(-\hat{H}(t))$
1	10	1	0	1/10	$9/10 = 0.9$	$1/10 = 0.1$	0.905
2	9	2	0	2/9	$9/10 \cdot (1-2/9) = 7/10 = 0.7$	$1/10 + 2/9 = 29/90 \approx 0.322$	0.725
4	7	0	1	0	0.7	0.322	0.725
5	6	0	1	0	0.7	0.322	0.725
6	5	1	0	1/5	$0.7 \cdot (1-1/5) = \frac{14}{25} = 0.56$	$\frac{29}{90} + \frac{1}{5} = \frac{47}{90} = 0.522$	0.593
7	4	0	1	0	0.56	0.522	0.593
8	3	0	1	0	0.56	0.522	0.593
9	2	0	1	0	0.56	0.522	0.593
10	1	0	1	0	0.56	0.522	0.593

P8131 HW9

3. Tongue data

Data Description:

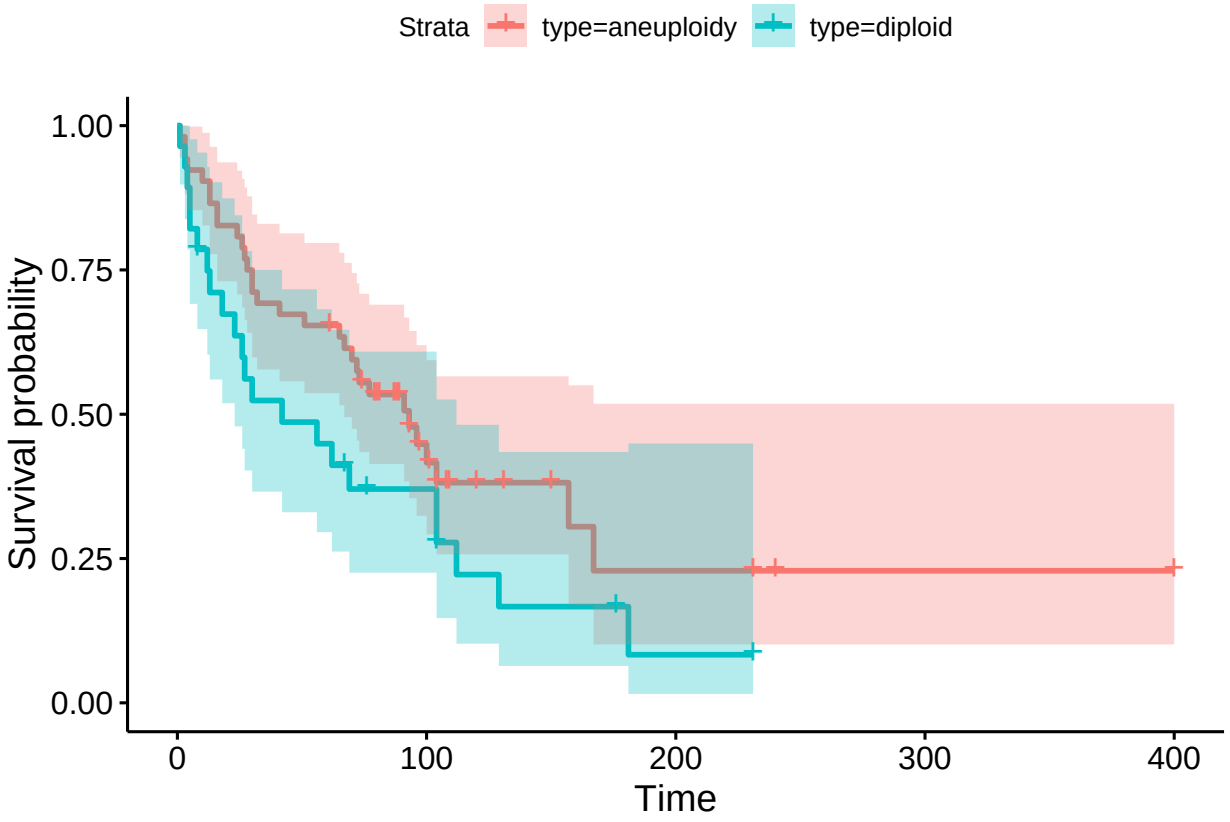
- `type`: Tumor DNA profile (1=Aneuploid Tumor, 2=Diploid Tumor)
- `time`: Time to death or on-study time, weeks
- `delta` Death indicator (0: censored/alive, 1: observed/dead)

```
data(tongue)
tongue <- tongue %>%
  mutate(type = ifelse(type == 1, "aneuploidy", "diploid"))
Surv(tongue$time,tongue$delta,type='right') # 0: censored, 1: observed
```

```
##  [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+
```

For each tumor type (aneuploidy and diploid), plot the Kaplan-Meier curve of survival function and its pointwise 95% confidence intervals (using the log transformation).

```
#survdiff(Surv(time,cens)~treat, data=gehan) # log rank test
survival.fit <- survfit(Surv(time,delta)~type, data = tongue, conf.type='log')
ggsurvplot(survival.fit, conf.int=TRUE)
```



The estimated 1-year survival rate and 95% CI:

```
summary(survival.fit,time=c(365/7))
```

```
## Call: survfit(formula = Surv(time, delta) ~ type, data = tongue, conf.type = "log")
##
##                 type=aneuploidy
##         time       n.risk      n.event     survival      std.err lower 95% CI
##       52.143       34.000       18.000        0.654        0.066        0.537
## upper 95% CI
##        0.797
##
##                 type=diploid
##         time       n.risk      n.event     survival      std.err lower 95% CI
##      52.1429      13.0000      14.0000       0.4864       0.0961       0.3302
## upper 95% CI
##       0.7164
```

The estimated 1-year survival rate for patients with aneuploid tumor is about 0.654, with 95% CI $[0.537, 0.797]$, and the estimated 1-year survival rate for patients with diploid tumor is about 0.4864, with 95% CI $[0.3302, 0.7164]$.