Problem 1 Hospital length of stay

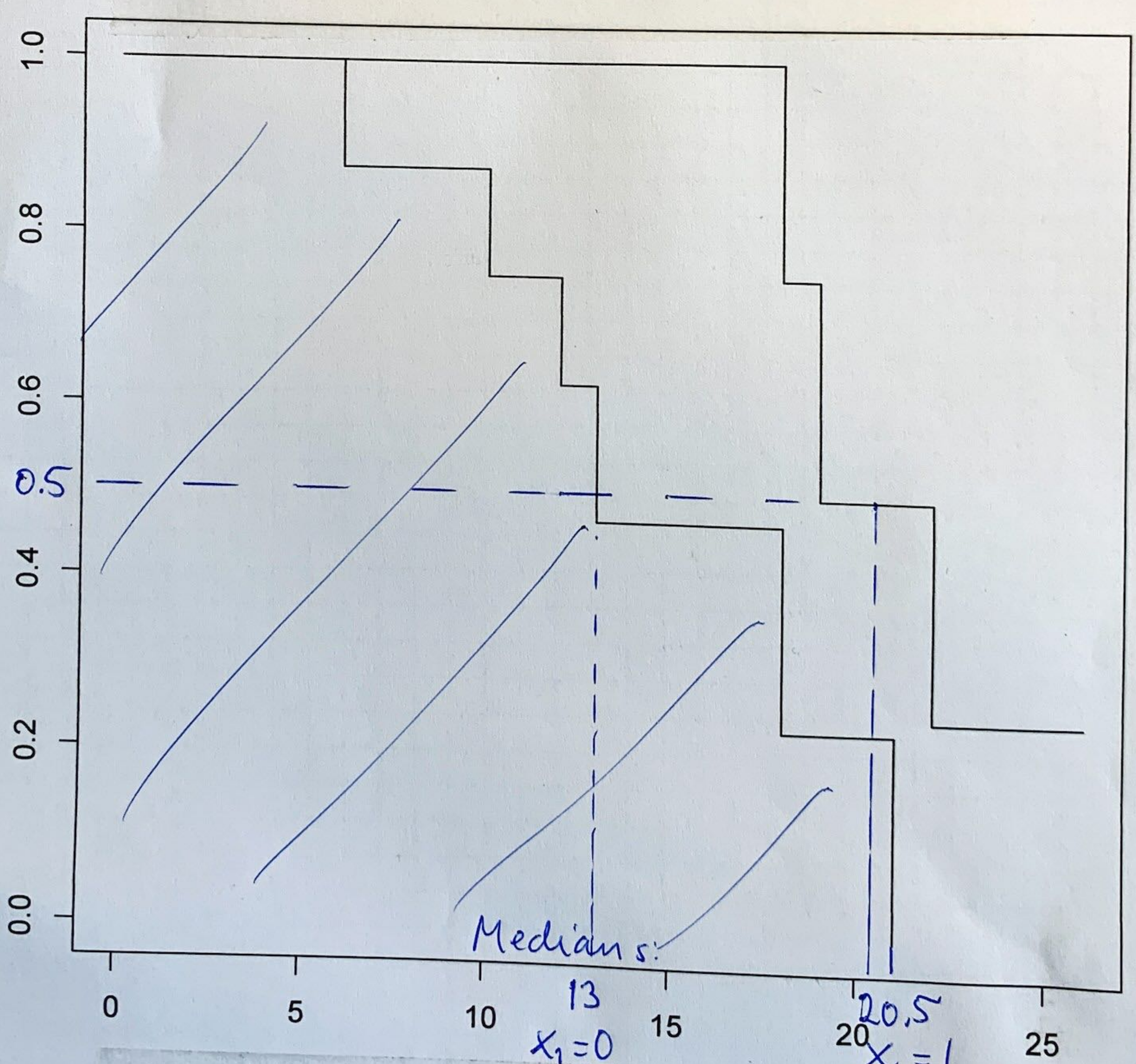
a)

strata(x1)=x1=0

time	n.risk	n.event	survival
6	8	1	0.875
10	7	1	0.750
12	6	1	0.625
13	4	1	0.469
18	2	1	0.234
21	1	1	0.000

strata(x1)=x1=1

time	n.risk	n.event	survival
18	4	1	0.75
19	3	1	0.50
22	2	1	0.25



Restricted mean length for first 21 days:
(See marked areas)

$x_1 = 0$: $6 \cdot 1 + 4 \cdot 0.875 + 2 \cdot 0.75 + 1 \cdot 0.625$
$+ 5 \cdot 0.469 + 3 \cdot 0.234 = 14.67$

$x_1 = 1$: $18 \cdot 1 + 1 \cdot 0.75 + 2 \cdot 0.5 = 19.75$

From R:

```
                 n events *rmean *se(rmean) median 0.95LCL 0.95UCL
strata(x1)=x1=0 9      6   14.7       1.86   13.0       6      NA
strata(x1)=x1=1 7      3   19.8       0.65   20.5      18      NA
    * restricted mean with upper limit =  21
>
```

A preliminary conclusion is that ~~a~~ increasing age may increase length of stay.

b) Relative risk function estimated:

$$e^{-0.44x_1 + 0.089x_2 + 0.0045x_3 - 0.25x_4 + 0.30x_5 - 0.46x_6}$$

Significant at 5%: $x1, x4, x_5, x_6$

Hazard ratio for age: (x_1) 0.644866

95% C.I.

$$e^{-0.438713 \pm 1.96 \cdot 0.120904}$$

gives $(0.509, 0.817)$

~~One year increase in age leads to multiplication~~

Age ≥ 45 years leads to multiplying the hazard for discharge by 0.645 as compared to age < 45

Hazard ratio for clinical grade (x_6): 0.629

A severe clinical grade leads to multiplying the hazard by $\underline{0.629}$.

Patients: (1) age ≥ 45, severe grade: $x_1 = 1, x_6 = 1$
(2) age < 45, mild grade: $x_1 = 0, x_6 = 0$

Relative risk:

$$e^{-0.44-0.46} = e^{-0.90} = \underline{0.407}$$

c)

1. One may add the covariate $x_1 \cdot x_6$ in the Cox analysis
   An interaction would be interpreted to say that, e.g., high age leads to a ~~higher~~ lower discharge hazard for severe patients, compared to low age for severe patients.
   In practice one may compare the partial likelihoods with and without the term $x_1 \cdot x_6$, and reject the null hypothesis of no interaction if $2 \cdot \text{difference} > \chi^2_{1,\alpha}$ (α-quantile in χ^2_1) for an α-level test

2. By using log cumulative hazards plot:

Can divide data into four groups:

(1) $x_1 = 0, x_6 = 0$

(2) $x_1 = 0, x_6 = 1$

(3) $x_1 = 1, x_6 = 0$

(4) $x_1 = 1, x_6 = 1$

For each group, estimate by Nelson-Aalen, plot their "logs". Should be parallell if model is correct.

By using Schoenfeld residuals: (see slides 12).

d) Model: $\alpha(t \mid x) = \beta_0(t) + \beta_1(t)x_1 + \beta_6(t)x_6$

Estimates $A(t \mid x) = B_0(t) + B_1(t)x_1 + B_6(t)x_6$

Intercept plot: Corresponds to the case when $x_1 = x_6 = 0$, so it is the cumulative hazard for age < 45, mild degree

The $\beta_0(t)$ is increasing with time (derivative of curve is increasing).

For $\beta_1(t)$ and $\beta_6(t)$ there is the same tendency that their effect on the hazard is decreasing with time, at least up to time 30.