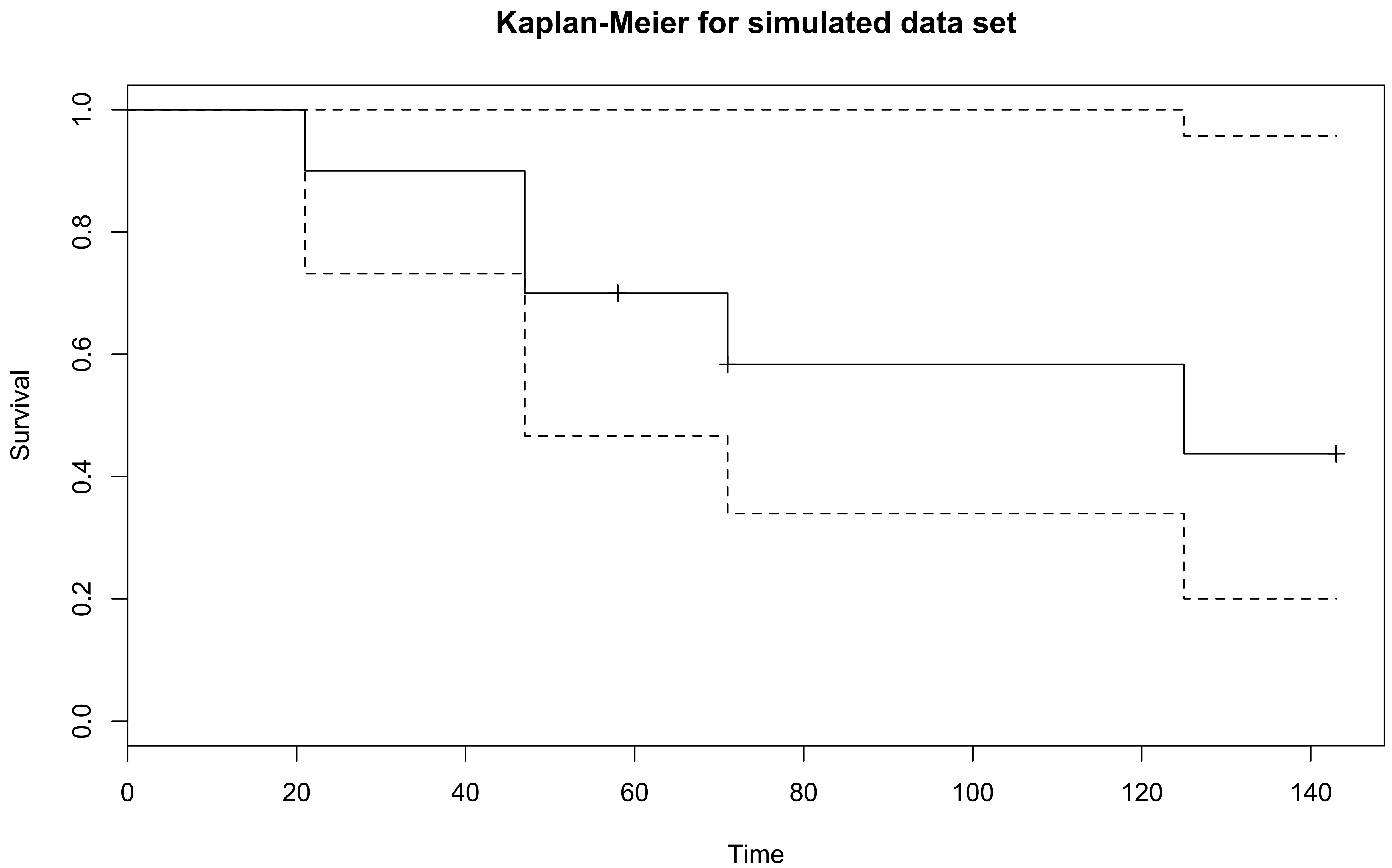
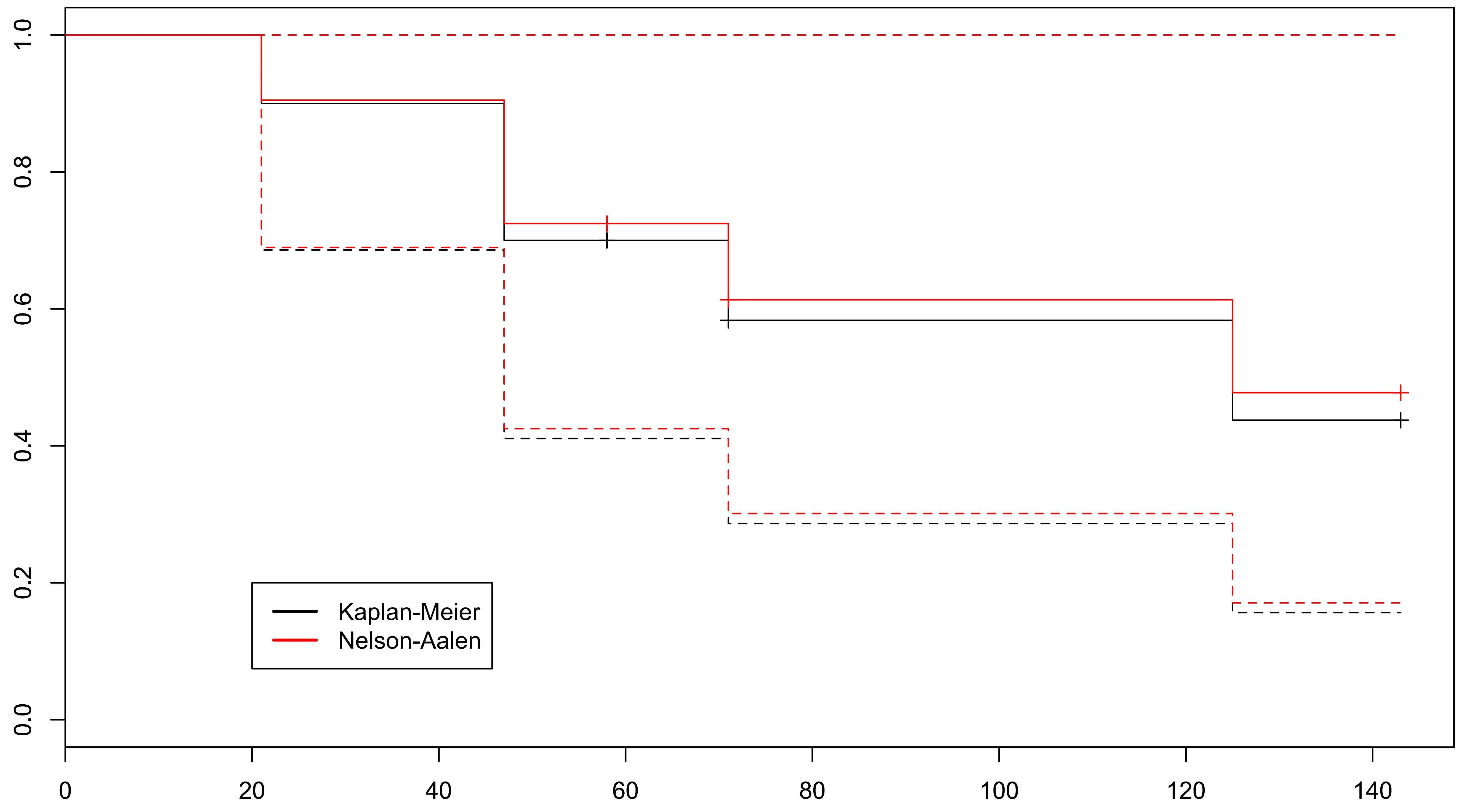
Estimation in R

```
require('survival')
sim.surv = Surv(c(21, 47, 47, 58, 71, 71, 125,
143, 143, 143), c(1, 1, 1, 0, 1, 0, 1, 0, 0, 0))

sim.fit=survfit(sim.surv~1,conf.int=.99)
plot(sim.fit)
```

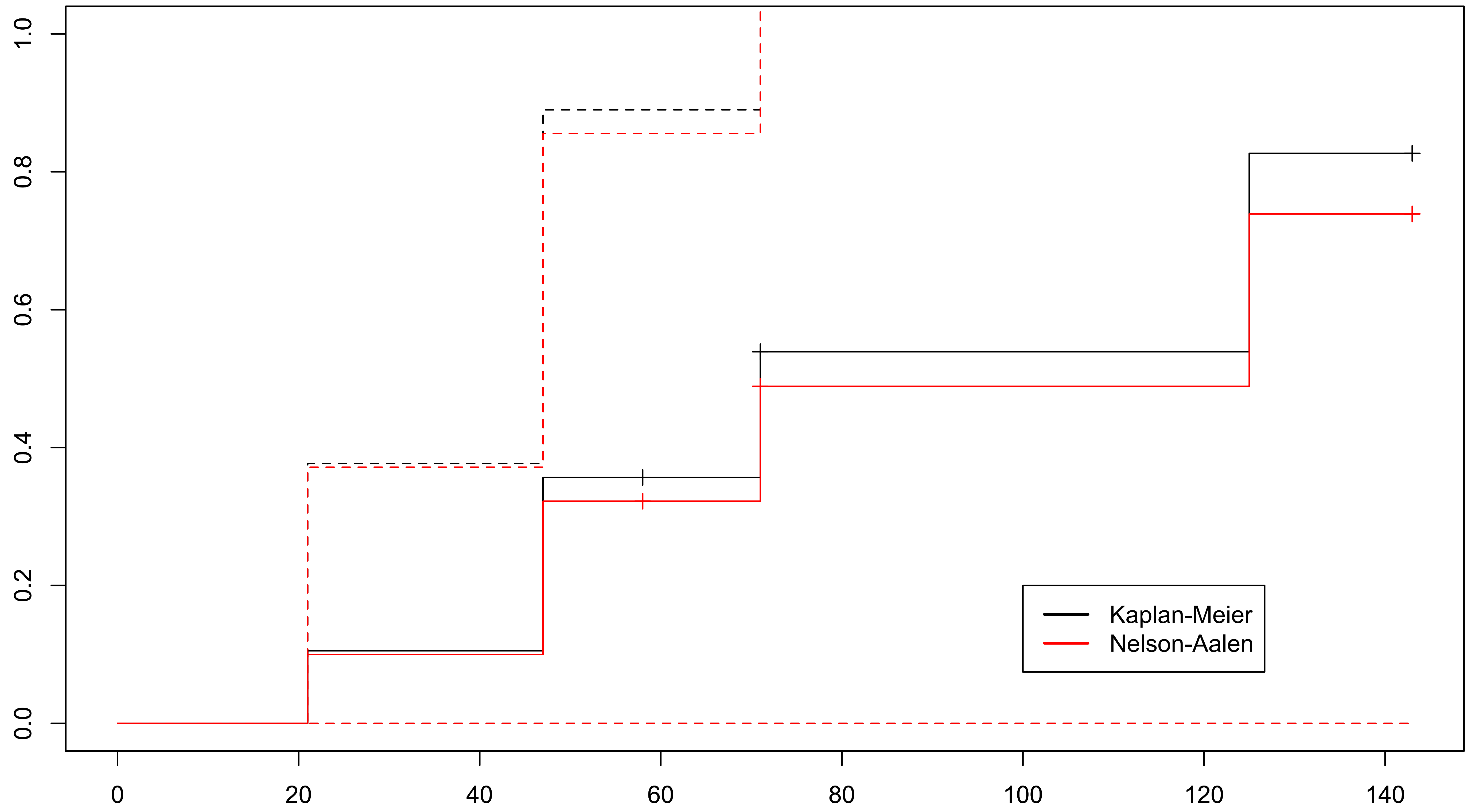


```
sim.fit2=survfit(sim.surv~1,conf.int=.99,type='fleming-
harrington')
par(new=TRUE)
plot(sim.fit2,col=2)

legend(20,.2,c('Kaplan-Meier','Nelson-Aalen'),col=1:2,lwd=2)
```



Cumulative hazard

```
plot(sim.fit,fun='cumhaz',ylim=c(0,1))
par(new=TRUE)
plot(sim.fit2,col=2,fun='cumhaz',ylim=c(0,1))

legend(100,.2,c('Kaplan-Meier','Nelson-Aalen'),col=1:2,lwd=2)
```



Testing for survival distributions

Single-sample Testing

period life data for e.g. a care home

age	μ^s_x	E_x (initial exposed to risk)	D_x	E_x^C	μ_x
90	0.202	40	10	35	0.29
91	0.215	35	8	31	0.258
92	0.236	22	4	18	0.20
93	0.261	14	6	11	0.545
94	0.279	11	4	9	0.444
95	0.291	7	3	5.5	0.545

Compare observed mortality to standard mortality rates μ_x^s

age	μ^s_x	E_x	D_x	E_x^c	μ_x
90	0.202	40	10	35	0.29
91	0.215	35	8	31	0.258
92	0.236	22	4	18	0.20
93	0.261	14	6	11	0.545
94	0.279	11	4	9	0.444
95	0.291	7	3	5.5	0.545

Each age class is an independent "experiment"

Continuous (Poisson) model

$$D_x \sim \mathcal{N}(E_x^c \mu_x, E_x^c \mu_x)$$

$$z_x = \frac{D_x - E_x^c \mu_x^s}{\sqrt{E_x^c \mu_x^s}} \quad \left(\approx \frac{O - E}{\sqrt{V}} \right).$$

Discrete (Binomial) model

$$D_x \sim B(E_x, q_x) \implies D_x \sim \mathcal{N}\left(E_x q_x, E_x q_x \left(1 - q_x\right)\right)$$

$$z_x = \frac{D_x - E_x q_x^s}{\sqrt{E_x q_x^s \left(1 - q_x^s\right)}} \quad \left(\approx \frac{O - E}{\sqrt{V}} \right).$$

Test the null hypothesis (data sampled from the same mortality distribution as the standard table) against the two-sided alternative (the mortality distributions are not the same)

Chi-squared test: $X = \sum_{\text{all ages } x} z_x^2 \sim \chi^2(m)$ for m age classes

Disadvantages:

- Won't detect isolated outliers.
- Underpowered for detecting bias in one direction.
- Ignores the order of age classes.

Cumulative deviations test

Continuous

$$Z := \frac{\sum \left(D_x - E_x^c \mu_x^s\right)}{\sqrt{\sum E_x^c \mu_x^s}} \sim \mathcal{N}(0,1) \text{ approximately.}$$

Discrete

$$Z := \frac{\sum \left(D_x - E_x q_x^s\right)}{\sqrt{\sum E_x q_x^s \left(1 - q_x^s\right)}} \sim N(0,1), \text{ approximately}$$

Better power for testing against biased alternative. Weak against “crossing” alternative.

Test can be one-sided or two-sided.

Other tests

- Signs test: $S = \#\{x : z_x > 0\} \sim \mathrm{Binom}(m, \frac{1}{2})$ under null hypothesis.
- Groups of signs test — detects groups of years with bias.
- Serial correlations test — also detects groups of years with bias.
- All need large number of years to obtain reasonable power.

age	μ^s_x	E_x^0	D_x	E_x^C	μ_x	z_x
90	0.202	40	10	35	0.29	1.1
91	0.215	35	8	31	0.258	0.52
92	0.236	22	4	18	0.20	-0.33
93	0.261	14	6	11	0.545	1.85
94	0.279	11	4	9	0.444	0.94
95	0.291	7	3	5.5	0.545	1.11

Continuous model

$$z_{93} = \frac{d_{93} - \mu^s_{93} E^c_{93}}{\sqrt{E^c_{93}\mu^s_{93}}}$$

$$= \frac{6 - 11 \cdot 0.261}{\sqrt{11 \cdot 0.261}} = \frac{3.129}{1.694} = 1.85.$$

Discrete model

$$z_{93} = \frac{d_{93} - q^s_{93} E^0_{93}}{\sqrt{E^0_{93} q^s_{93}(1 - q^s_{93})}}$$

$$= \frac{6 - .230 \cdot 14}{\sqrt{14 \cdot .230 \cdot .770}} = 1.77.$$

Chi-squared: $X^2 = \sum_{\text{all ages } x} z_x^2 = 7.13$

```
> 1-pchisq(7.13,6)
[1] 0.3089921
```

Signs test: $S = \# \{x : z_x > 0\} = 5$

```
> 2*pbinom(1,6,.5)
[1] 0.21875
```

Cumulative deviations test:

$$Z = \frac{\sum (D_x - E^C_x \mu^s_x)}{\sqrt{\sum E^C_x \mu^s_x}} = 2.01$$

```
> 2*(1-pnorm(2.01))
[1] 0.04443119
```

Graduation

age	l_x	D_x	E_x^c	μ_x	μ_x^s	graduated $\mathring{\mu}_x$	z_x
90	40	10	35	0.29	0.202	0.25	0.42
91	35	8	31	0.258	0.215	0.28	-0.23
92	22	4	18	0.20	0.236	0.335	-0.83
93	14	6	11	0.545	0.261	0.40	0.76
94	11	4	9	0.444	0.279	0.45	-0.02
95	7	3	5.5	0.545	0.291	0.48	0.22

$$\mathring{\mu}_x = -0.279 + 2.6\mu_x^s$$

$$X = \sum z_x^2 = 1.55$$

p-value

```
1-pchisq(1.55,4)=0.818
```

Two-sample testing for survival distributions

Event times are $0 < t_1 < t_2 < \cdots < t_m$.

For $j = 1,2,\ldots,m,$ **and** $i = 1,2,$ $d_{ij} = \#$ **events at** t_j **in group** i,

$n_{ij} = \#$ **in risk set at** t_j **from group** i,

$d_j = \#$ **events at** t_j,

$n_j = \#$ **in risk set at** t_j.

Test null hypothesis H_0: Both samples came from populations with the same hazard against the two-sided alternative that the hazard rates differ.

No ties: Probability (under H_0) that the event at time t_j is from group *i* is n_{ij}/n_j.

More generally, under H_0 conditioned on d_j events at time t_j the number of events from group *i* has a hypergeometric distribution with

$$\text{expectation} = d_j \frac{n_{1j}}{n_j}, \text{ and}$$

$$\text{variance} =: \sigma_j^2 = \frac{n_{1j} n_{2j} (n_j - d_j) d_j}{n_j^2 (n_j - 1)}.$$

Two-sample testing for survival distributions

Test null hypothesis H_0: Both samples came from populations with the same hazard against the two-sided alternative that the hazard rates differ.

More generally, under H_0 conditioned on d_j events at time t_j the number of events from group *i* has a hypergeometric distribution with

$$\text{expectation} = d_j \frac{n_{1j}}{n_j}, \text{ and}$$

$$\text{variance} =: \sigma_j^2 = \frac{n_{1j} n_{2j} (n_j - d_j) d_j}{n_j^2 (n_j - 1)}.$$

$$M := \sum_{j=1}^{m} W(t_j) \left(d_{1j} - n_{1j} \frac{d_j}{n_j} \right)$$

is a random variables with expectation 0 and variance $\sum_{i=1}^{k} W(t_j)^2 \sigma_j^2$

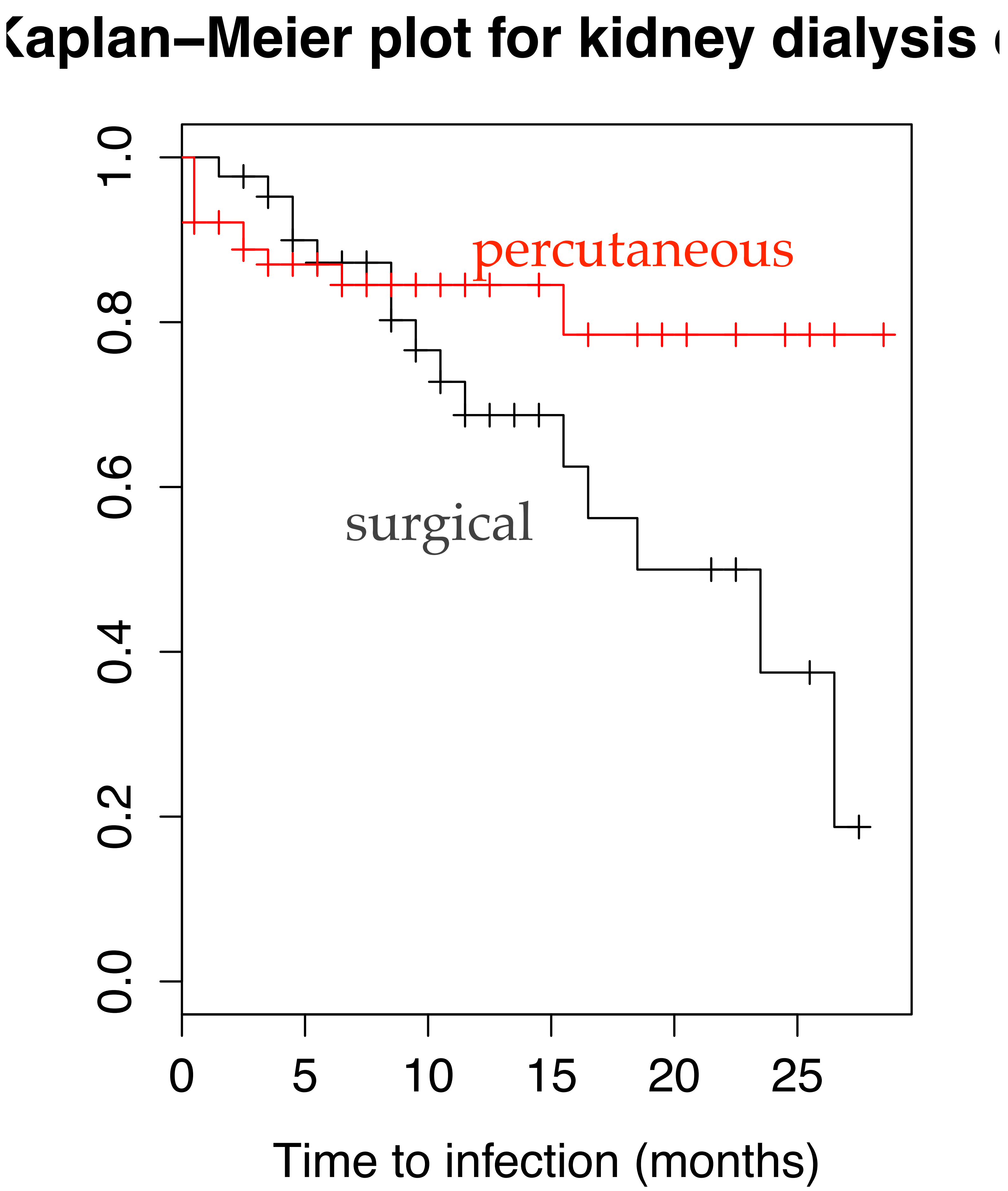
$$Z := \frac{\sum_{j=1}^{m} W(t_j) \left(d_{1j} - n_{1j} \frac{d_j}{n_j} \right)}{\sqrt{\sum_{j=1}^{m} W(t_j)^2 \frac{n_{1j} n_{2j} (n_j - d_j) d_j}{n_j^2 (n_j - 1)}}}$$

is an approximately standard normal test statistic under the null hypothesis.

Standard tests

- Log-rank test: $W(t_j) = 1$
- Petos' test $W(t_j) = \widetilde{S}(t_{i-1})\dfrac{n_j}{n_j + 1}$, where $\widetilde{S}$ is a modified survival estimate $\widetilde{S}(t) = \prod_{t_j \le t} \dfrac{n_j + 1 - d_j}{n_j + 1}$
- Gehan, Breslow: $W(t_j) = n_j$.
- Fleming—Harrington tests: $W(t_j) = \left(\widehat{S}(t_{j-1})\right)^p \left(1 - \widehat{S}(t_{j-1})\right)^q$.
    - (1,0) close to Petos' test
    - (0,0) is log-rank test

Kidney dialysis example



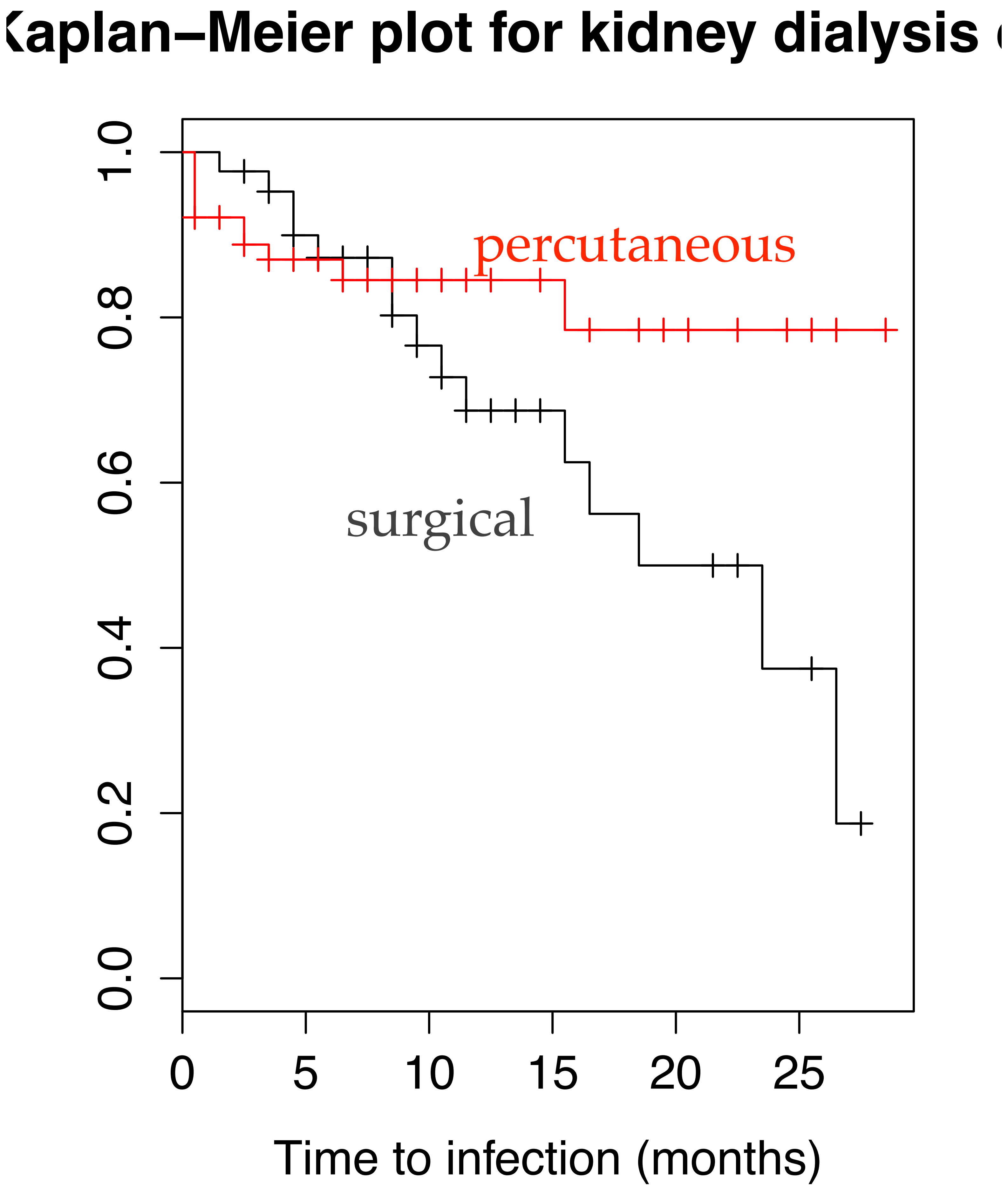
t_j	n_{1j}	n_{2j}	d_{1j}	d_{2j}	σ_j^2	Peto wt.	H–F (0, 1) wt.
0.5	43	76	0	6	1.326	0.992	0.000
1.5	43	60	1	0	0.243	0.941	0.050
2.5	42	56	0	2	0.485	0.931	0.059
3.5	40	49	1	1	0.489	0.912	0.078
4.5	36	43	2	0	0.490	0.890	0.099
5.5	33	40	1	0	0.248	0.867	0.121
6.5	31	35	0	1	0.249	0.854	0.133
8.5	25	30	2	0	0.487	0.839	0.146
9.5	22	27	1	0	0.247	0.807	0.176
10.5	20	25	1	0	0.247	0.790	0.193
11.5	18	22	1	0	0.247	0.770	0.210
15.5	11	14	1	1	0.472	0.741	0.230
16.5	10	13	1	0	0.246	0.681	0.289
18.5	9	11	1	0	0.247	0.649	0.319
23.5	4	3	1	0	0.245	0.568	0.351
26.5	2	3	1	0	0.240	0.473	0.432

Z_{LR}=1.59

Z_{Peto}=1.12

Z_{HF}=3.11

Kidney dialysis example




```
> kS=Surv(kidney$time,kidney$delta)
> survdiff(kS~kidney$type) #log-rank test
Call:
survdiff(formula = kS ~ kidney$type)

                N Observed Expected (O-E)^2/E (O-E)^2/V
kidney$type=1 43       15       11      1.42      2.53
kidney$type=2 76       11       15      1.05      2.53

 Chisq= 2.5  on 1 degrees of freedom, p= 0.112
> survdiff(kS~kidney$type,rho=1) #H-F(1,0) test
Call:
survdiff(formula = kS ~ kidney$type, rho = 1)

                N Observed Expected (O-E)^2/E (O-E)^2/V
kidney$type=1 43     12.0     9.48     0.686      1.39
kidney$type=2 76     10.4    12.98     0.501      1.39

 Chisq= 1.4  on 1 degrees of freedom, p= 0.239
```

$Z_{LR}=1.59$

$Z_{Peto}=1.12$

$Z_{HF(1,0)}=1.18=\text{sqrt}(1.39)$