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

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

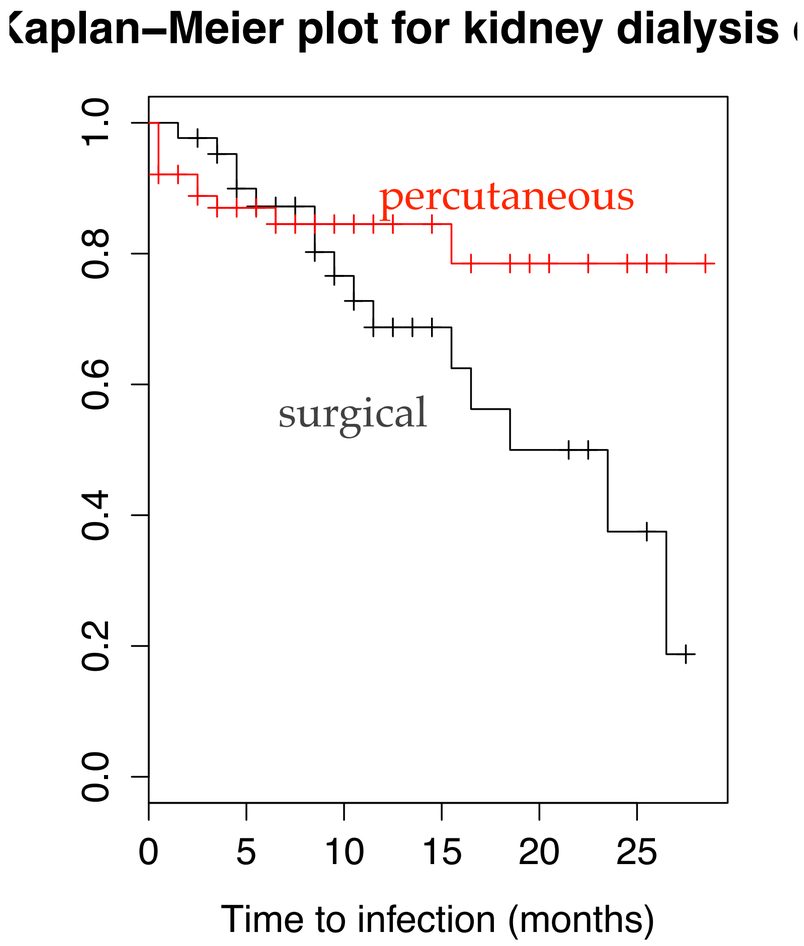
$$\text{variance} =: \sigma_j^2 = \frac{n_{1j} n_{2j} (n_j - d_j) d_j}{n_j^2 (n_j - 1)}.$$

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

Kidney dialysis example



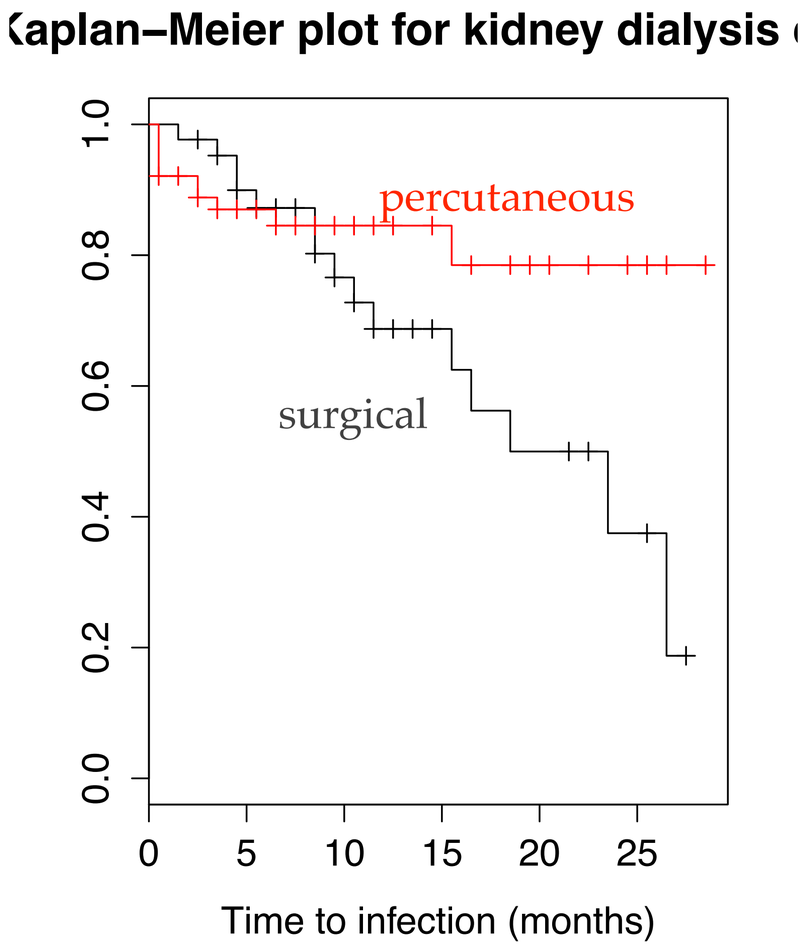
t_j	n_{1j}	n_{2j}	d_{1j}	d_{2j}	σ_j^2	Peto wt.	H–F $(0,1)$ wt.
0.5	43	76	0	6	1.326	0.992	0.000
1.5	43	60	1	0	0.243	0.941	0.050
2.5	42	56	0	2	0.485	0.931	0.059
3.5	40	49	1	1	0.489	0.912	0.078
4.5	36	43	2	0	0.490	0.890	0.099
5.5	33	40	1	0	0.248	0.867	0.121
6.5	31	35	0	1	0.249	0.854	0.133
8.5	25	30	2	0	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
```

$Z_{LR}=1.59$

$Z_{Peto}=1.12$

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

Regression in a survival context (LN chap. 5)

Types of regression models

- **Proportional hazards** (or **relative risk**): Standard hazard *h₀(t)* at time *t*. Individual *i* has hazard $h_i(t) = \rho_i h_0(t)$ at time *t*. Equivalently, $S_i(t) = S_0(t)^{\rho_i}$.

- **Accelerated lifetime** (or **accelerated failure**): $S_i(t) = S_0(\rho_i t)$. Equivalently, $h_i(t) = \rho_i h_0(\rho_i t)$.

- **Additive hazards**: $h_i(t) = h_0(t) + \rho_i(t)$. Here ρ_i is a function, and so needs further specification to be estimated.

Generalised linear survival models

Observe for each individual

- Response
  - Event time T_i ;
  - Status δ_i (=1 if failure, =0 if censored);
  - Possibly a left censoring time.
- Covariates $x_i(t)$ such as
  - Age at entry;
  - Sex;
  - Blood pressure;
  - Treatment group.

Linear risk score $\beta \cdot \mathbf{x}_i$

Risk $\rho_i = \rho(\beta \cdot \mathbf{x}_i) = \psi^{-1}(\beta \cdot \mathbf{x}_i)$ where ψ is the **link.**

Most common log link, so $\rho(\beta \cdot \mathbf{x}) = \mathrm{e}^{\beta \cdot \mathbf{x}}$.

Example: Overall survival distribution Weibull

$$S_i(t) = \mathrm{e}^{-(\rho_i t)^{\alpha}} \text{ and } \rho_i = \mathrm{e}^{\beta \cdot \mathbf{x}_i}$$
$$\beta \cdot \mathbf{x} = \beta_0 + \beta_1 \mathrm{age}_i + \beta_2 \mathrm{sex}_i + \beta_3 \mathrm{sbp}_i + \beta_4 \mathrm{trt}_i \,.$$

This is AL and PH.

Likelihood $L(\alpha, \beta) = \prod_i \left(\alpha \rho_i^{\alpha} t_i^{\alpha-1}\right)^{\delta_i} \mathrm{e}^{-(\rho_i t_i)^{\alpha}} = \prod_i \left(\alpha \mathrm{e}^{\alpha\beta \cdot \mathbf{x}^i} t_i^{\alpha-1}\right)^{\delta_i} \mathrm{e}^{-(\mathrm{e}^{\beta \cdot x_i} t_i)^{\alpha}}$.

Now compute MLE for all components, and we compute standard errors for α and β from Fisher information.

Generalised linear survival models

Linear risk score $\beta \cdot \mathbf{x}_i$

Risk $\rho_i = \rho(\beta \cdot \mathbf{x}_i) = \psi^{-1}(\beta \cdot \mathbf{x}_i)$ where ψ is the **link**.

Most common log link, so $\rho(\beta \cdot \mathbf{x}) = \mathrm{e}^{\beta \cdot \mathbf{x}}$.

Example: Overall survival distribution Weibull

$$S_i(t) = \mathrm{e}^{-(\rho_i t)^\alpha} \text{ and } \rho_i = \mathrm{e}^{\beta \cdot \mathbf{x}_i}$$
$$\beta \cdot \mathbf{x} = \beta_0 + \beta_1 \mathrm{age}_i + \beta_2 \mathrm{sex}_i + \beta_3 \mathrm{sbp}_i + \beta_4 \mathrm{trt}_i \,.$$

This is AL and PH.

Likelihood
$$L(\alpha, \beta) = \prod_i \left(\alpha \rho_i^\alpha t_i^{\alpha-1}\right)^{\delta_i} \mathrm{e}^{-(\rho_i t_i)^\alpha}$$
$$= \prod_i \left(\alpha \mathrm{e}^{\alpha\beta \cdot \mathbf{x}^i} t_i^{\alpha-1}\right)^{\delta_i} \mathrm{e}^{-\left(\mathrm{e}^{\beta \cdot x_i} t_i\right)^\alpha} .$$

More generally, if we have any parametric survival model with cumulative hazard $H_\alpha(t)$ (α represents parameters that do not vary) and hazard $h_\alpha(t) = H'_\alpha(t)$. Then we can make an AL model with the likelihood

$$L(\alpha, \beta) = \prod_i \left[\mathrm{e}^{\beta \cdot \mathbf{x}^i} h_\alpha\left(\mathrm{e}^{\beta \cdot \mathbf{x}^i} t_i\right)\right]^{\delta_i} \mathrm{e}^{-H_\alpha(\mathrm{e}^{\beta \cdot \mathbf{x}^i} t_i)}$$

and PH model with the likelihood

$$L(\alpha, \beta) = \prod_i \left[\mathrm{e}^{\beta \cdot \mathbf{x}^i} h_\alpha\left(t_i\right)\right]^{\delta_i} \mathrm{e}^{-\mathrm{e}^{\beta \cdot \mathbf{x}^i} H_\alpha(t_i)} .$$

Graphical tests

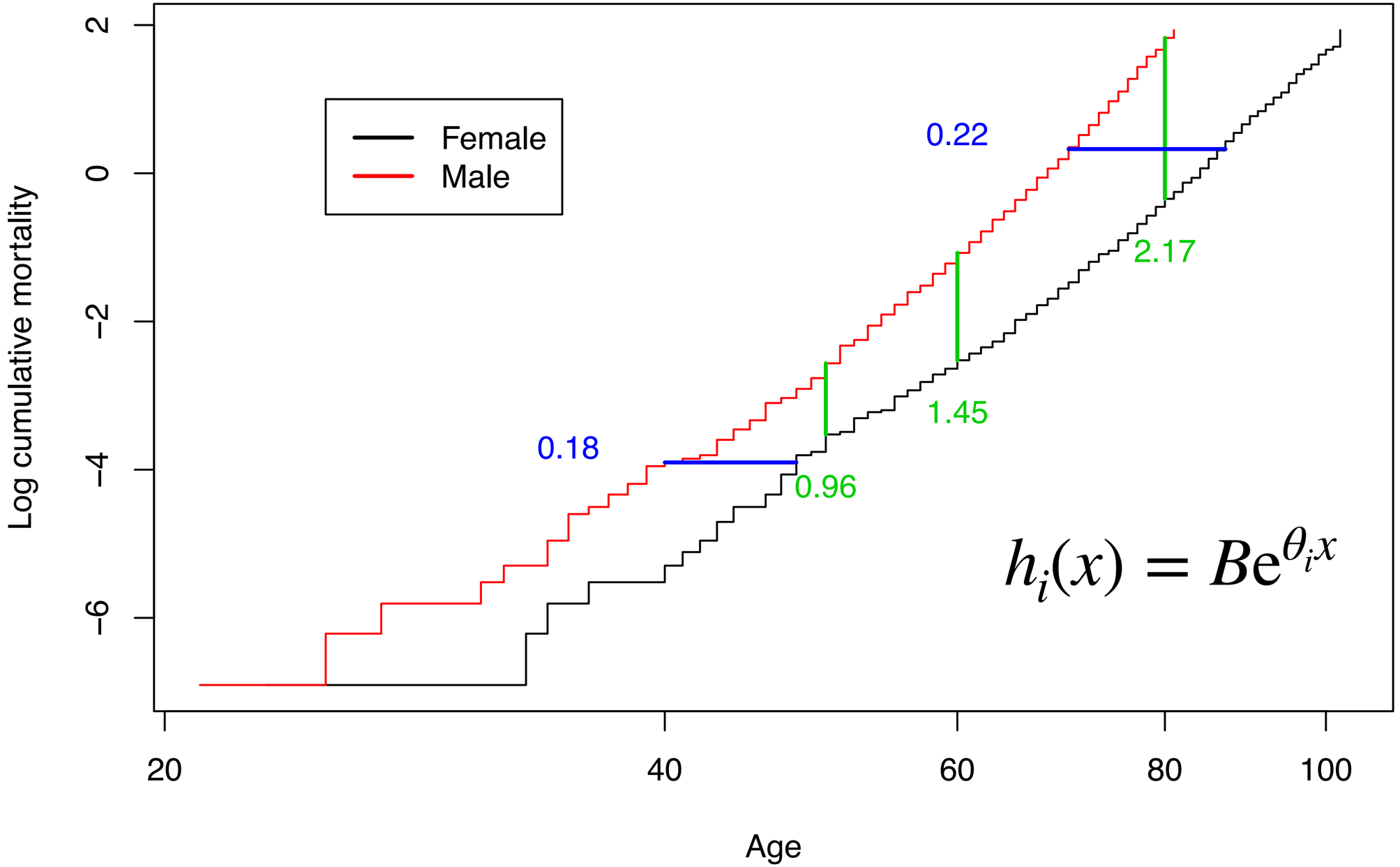
- Given categorical covariates with small number of categories, we make specific plots to test whether AL or PH is reasonable. Suppose we have groups g=1,2, …,*k.*

- Test AL assumption: $S_g(t) = S_0(\rho_g t) = S_0(\mathrm{e}^{\log \rho_g + \log t})$. So if we plot estimated survival probabilities for different groups against log*t,* the curves should differ by a horizontal shift.

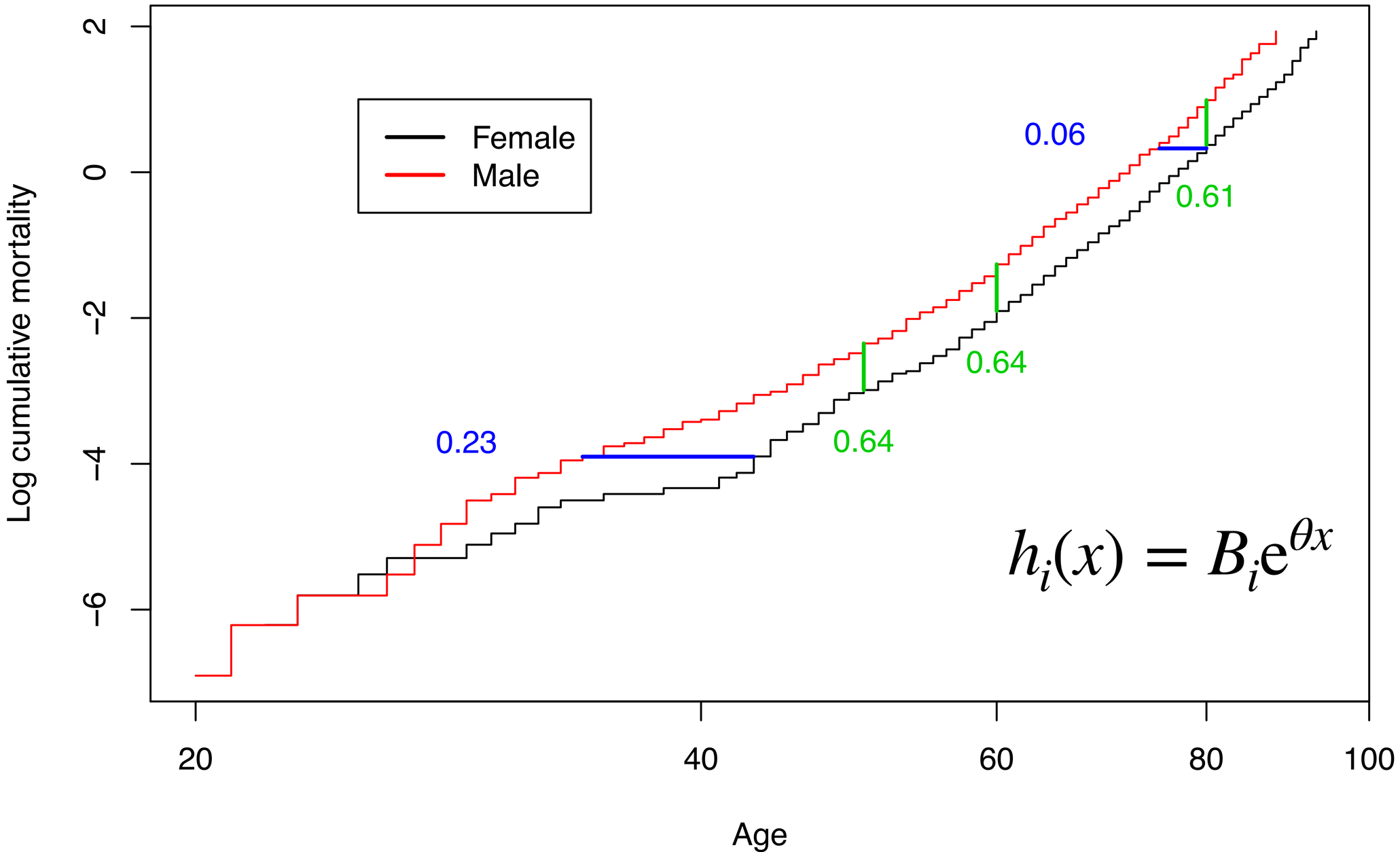
- Test PH assumption: $\log H_g(t) = \log \rho_g + \log H_0(t)$. So if we plot log estimated cumulative hazard for different groups against *t,* the curves should differ by a vertical shift.

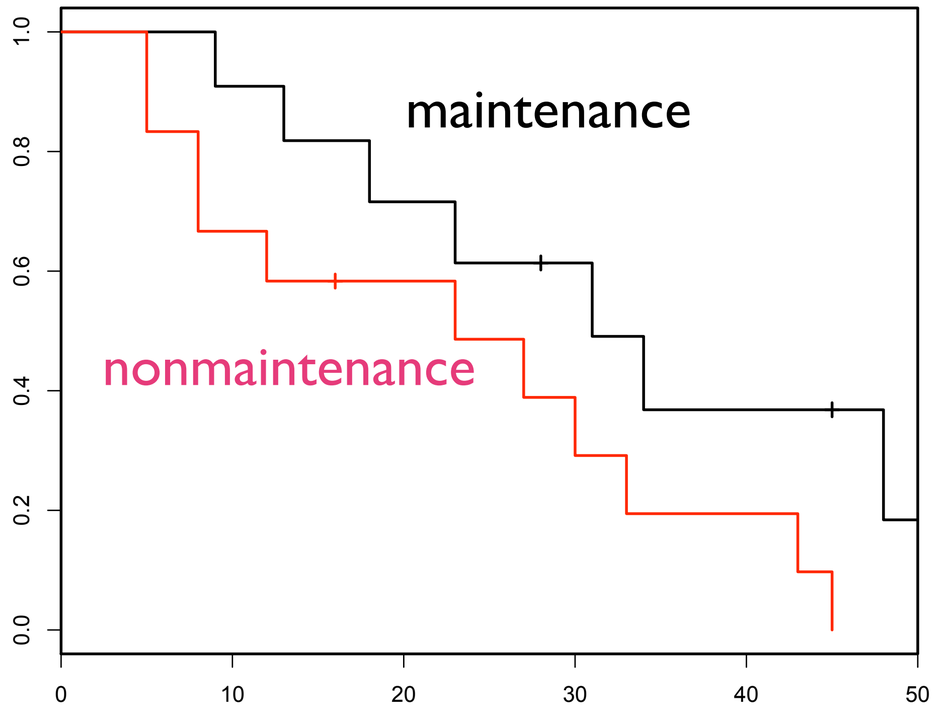
- Test both: Plot log estimated cumulative hazard against log *t.*

Accelerated lifetime plot



Proportional Hazards

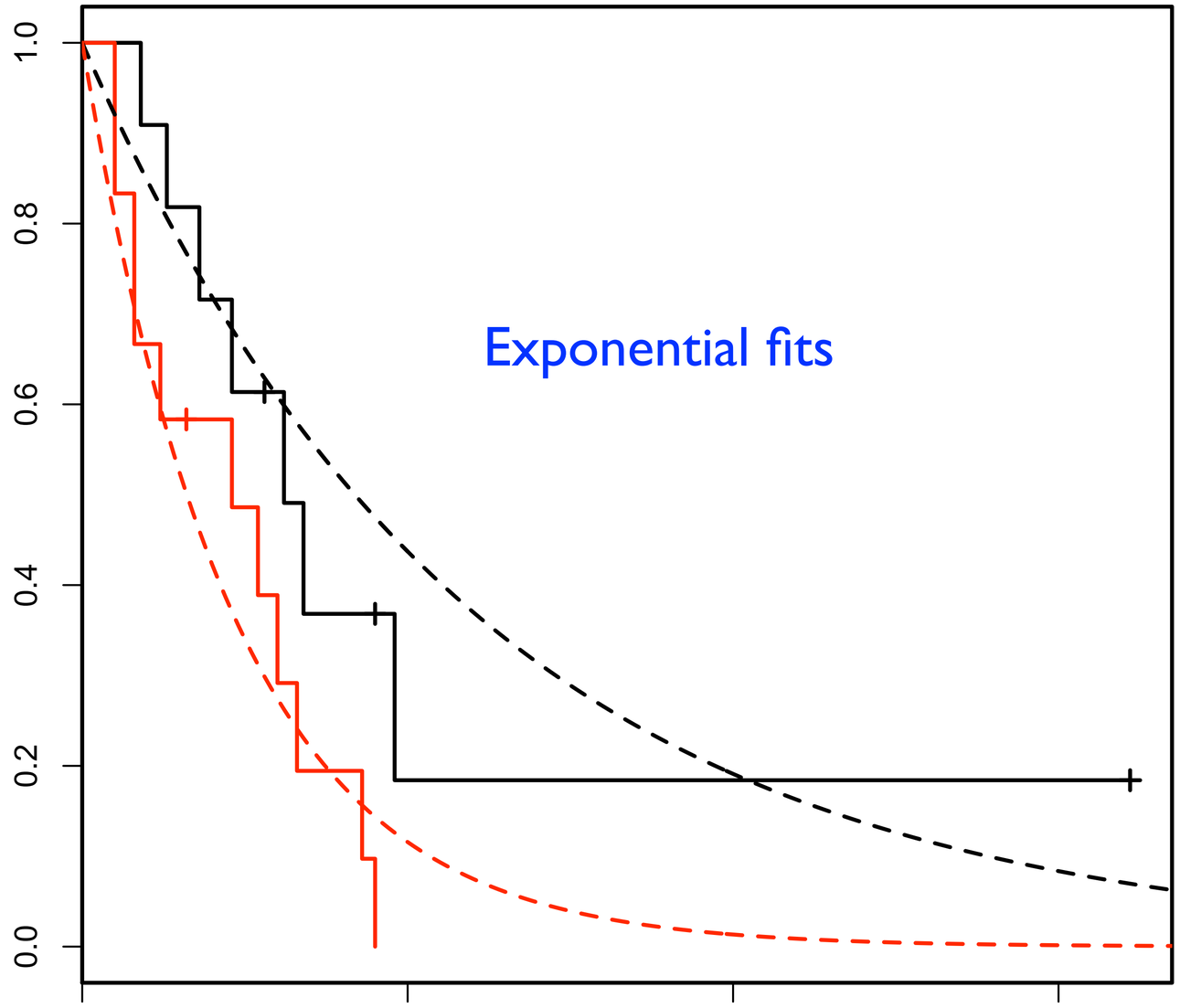




	Maintenance						Non-Maintenance (control)					
t_i	n_i	d_i	$\hat{h}_i$	$\hat{S}(t_i)$	$\hat{H}_i$	$\widetilde{S}(t_i)$	n_i	d_i	$\hat{h}_i$	$\hat{S}(t_i)$	$\hat{H}_i$	$\widetilde{S}(t_i)$
5	11	0	0.00	1.00	0.00	1.00	12	2	0.17	0.83	0.17	0.85
8	11	0	0.00	1.00	0.00	1.00	10	2	0.20	0.67	0.37	0.69
9	11	1	0.09	0.91	0.09	0.91	8	0	0.00	0.67	0.37	0.69
12	10	0	0.00	0.91	0.09	0.91	8	1	0.12	0.58	0.49	0.61
13	10	1	0.10	0.82	0.19	0.83	7	0	0.00	0.58	0.49	0.61
18	8	1	0.12	0.72	0.32	0.73	6	0	0.00	0.58	0.49	0.61
23	7	1	0.14	0.61	0.46	0.63	6	1	0.17	0.49	0.66	0.52
27	6	0	0.00	0.61	0.46	0.63	5	1	0.20	0.39	0.86	0.42
30	5	0	0.00	0.61	0.46	0.63	4	1	0.25	0.29	1.11	0.33
31	5	1	0.20	0.49	0.66	0.52	3	0	0.00	0.29	1.11	0.33
33	4	0	0.00	0.49	0.66	0.52	3	1	0.33	0.19	1.44	0.24
34	4	1	0.25	0.37	0.91	0.40	2	0	0.00	0.19	1.44	0.24
43	3	0	0.00	0.37	0.91	0.40	2	1	0.50	0.10	1.94	0.14
45	3	0	0.00	0.37	0.91	0.40	1	1	1.00	0.00	2.94	0.05
48	2	1	0.50	0.18	1.41	0.24	0	0				

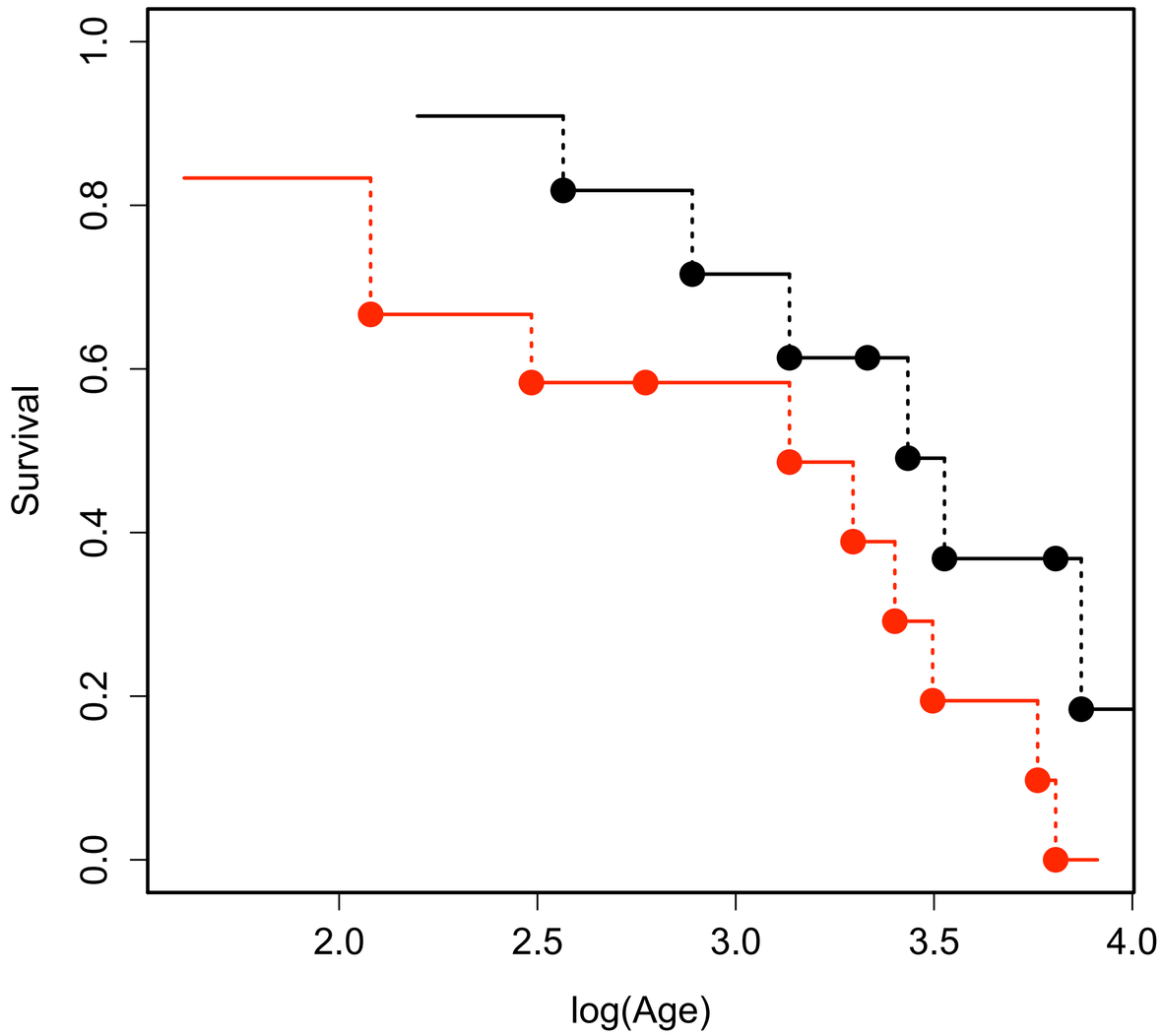
Maintained: 9 13 13+ 18 23 28+ 31 34 45+ 48 161+

Nonmaintained: 5 5 8 8 12 16+ 23 27 30 33 43 45



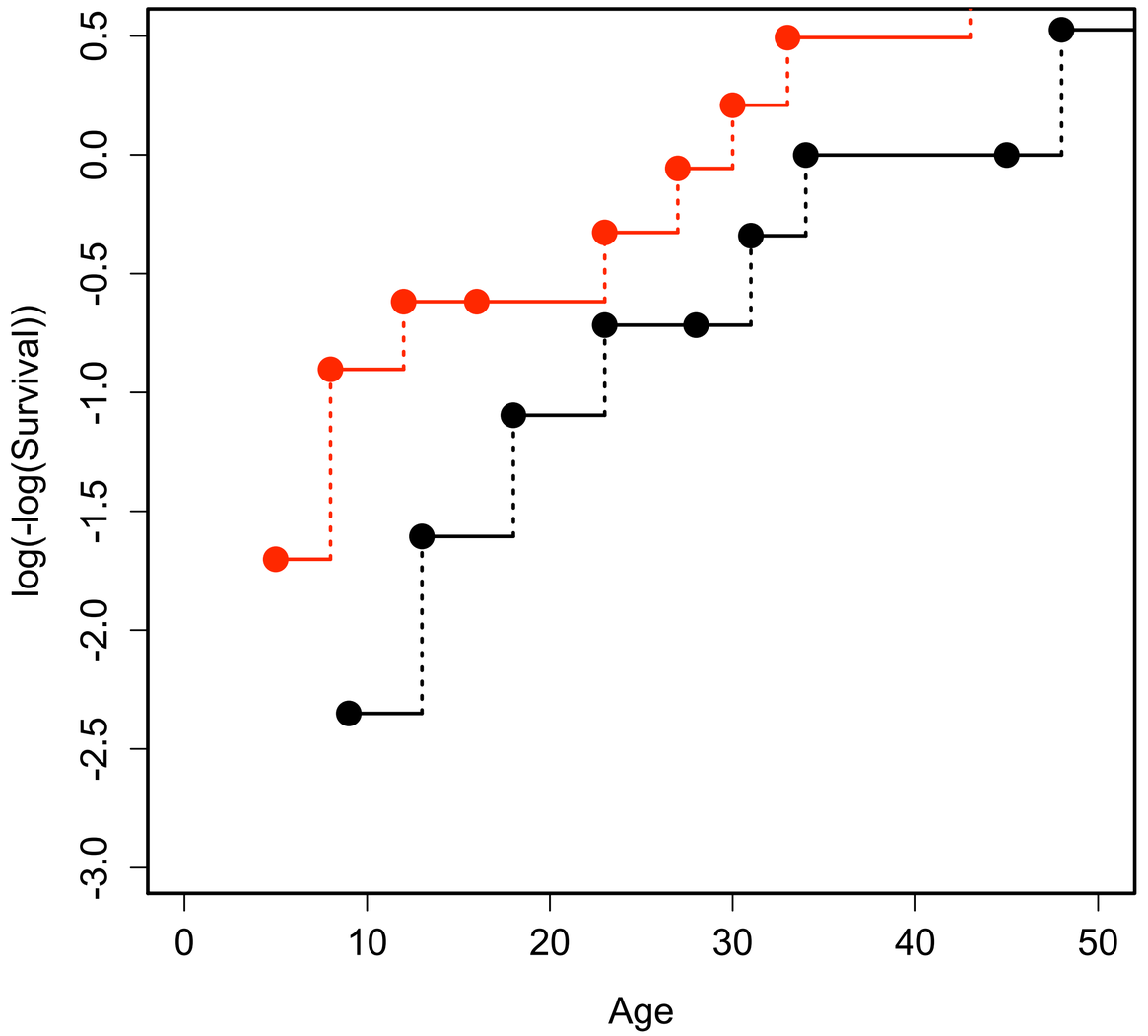
Maintained: 9 13 13+ 18 23 28+ 31 34 45+ 48 161+

Nonmaintained: 5 5 8 8 12 16+ 23 27 30 33 43 45



Maintained: 9 13 13+ 18 23 28+ 31 34 45+ 48 161+

Nonmaintained: 5 5 8 8 12 16+ 23 27 30 33 43 45



Maintained: 9 13 13+ 18 23 28+ 31 34 45+ 48 161+

Nonmaintained: 5 5 8 8 12 16+ 23 27 30 33 43 45