

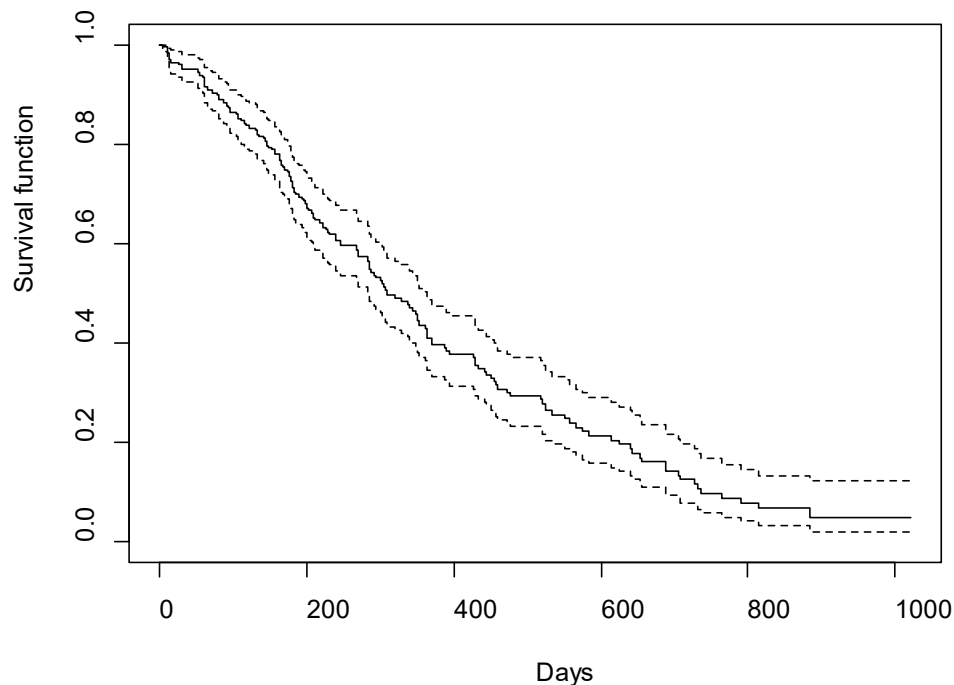
Survival Analysis

Kaplan-Meier estimation. Log-rank test. Cox proportional hazards model.

```
> library(survival)
> head(lung)      # Lung cancer survival data from a clinical trial, already in R
  inst time status age sex ph.ecog ph.karno pat.karno meal.cal wt.loss
1     3  306      2  74  1         1         90         100      1175      NA
2     3  455      2  68  1         0         90          90      1225      15
3     3 1010      1  56  1         0         90          90        NA      15
4     5  210      2  57  1         1         90          60      1150      11
5     1  883      2  60  1         0        100          90        NA         0
6    12 1022      1  74  1         1         50          80       513         0
> ?lung           # To learn more about the lung data
```

Kaplan-Meier estimation of the survival function:

```
> plot(survfit( Surv(time,status) ~ 1, data=lung ), xlab="Days",ylab="Survival function")
```



To see the details and the estimated survival probabilities, save the fit:

```
> fit = survfit( Surv(time,status) ~ 1, data=lung )
> summary(fit)
```

| time | n.risk | n.event | survival | std.err | lower 95% CI | upper 95% CI |
|----------------|--------|---------|----------|---------|--------------|--------------|
| 5 | 228 | 1 | 0.9956 | 0.00438 | 0.9871 | 1.000 |
| 11 | 227 | 3 | 0.9825 | 0.00869 | 0.9656 | 1.000 |
| 12 | 224 | 1 | 0.9781 | 0.00970 | 0.9592 | 0.997 |
| 13 | 223 | 2 | 0.9693 | 0.01142 | 0.9472 | 0.992 |
| Etc. etc. etc. | | | | | | |
| 814 | 7 | 1 | 0.0671 | 0.02351 | 0.0338 | 0.133 |
| 883 | 4 | 1 | 0.0503 | 0.02285 | 0.0207 | 0.123 |

Log-rank test

Here we compare two survival curves, determined by gender.
 # Option rho=0 calls the log-rank test

```
> survdiff( Surv(time,status) ~ sex, rho=0, data=lung )
```

| | N | Observed | Expected | (O-E)^2/E | (O-E)^2/V |
|-------|-----|----------|----------|-----------|-----------|
| sex=1 | 138 | 112 | 91.6 | 4.55 | 10.3 |
| sex=2 | 90 | 53 | 73.4 | 5.68 | 10.3 |

Chisq= 10.3 on 1 degrees of freedom, p= 0.001

Fit the Cox proportional hazards model

```
> coxph( Surv(time,status) ~ age + sex + meal.cal + wt.loss, data=lung )
```

| | coef | exp(coef) | se(coef) | z | p |
|----------|------------|-----------|-----------|--------|--------|
| age | 0.0178260 | 1.0179858 | 0.0110505 | 1.613 | 0.1067 |
| sex | -0.4638206 | 0.6288764 | 0.1975423 | -2.348 | 0.0189 |
| meal.cal | -0.0001201 | 0.9998799 | 0.0002469 | -0.486 | 0.6267 |
| wt.loss | -0.0005425 | 0.9994576 | 0.0067778 | -0.080 | 0.9362 |

Likelihood ratio test=10.07 on 4 df, p=0.03919

Output: coefficients β , their exponents e^β = IRR, the standard errors of estimated coefficients, the z-scores, and the p-values.

For confidence intervals and more details, do the summary:

```
> Cox = coxph(Surv(time,status) ~ age+sex+meal.cal+wt.loss, data=lung)
> summary(Cox)
```

| | exp(coef) | exp(-coef) | lower .95 | upper .95 |
|----------|-----------|------------|-----------|-----------|
| age | 1.0180 | 0.9823 | 0.9962 | 1.0403 |
| sex | 0.6289 | 1.5901 | 0.4270 | 0.9262 |
| meal.cal | 0.9999 | 1.0001 | 0.9994 | 1.0004 |
| wt.loss | 0.9995 | 1.0005 | 0.9863 | 1.0128 |

Likelihood ratio test= 10.07 on 4 df, p=0.04

Wald test = 9.63 on 4 df, p=0.05

Score (logrank) test = 9.78 on 4 df, p=0.04