Parametric Survival Models

Md. Moyazzem Hossain
Professor,
Dept. of Statistics & Data Science
Jahangirnagar University

Parametric Survival Models

Parametric survival models are a class of models in survival analysis that make explicit assumptions about the distribution of survival times.

Unlike non-parametric methods (e.g., Kaplan-Meier estimator) or semi-parametric methods (e.g., Cox Proportional Hazards model), parametric models specify a functional form for the survival distribution.

Parametric Survival Models

This allows for more specific modeling of the underlying survival time distribution.

Commonly used parametric survival models include

- Exponential,
- Weibull,
- Log-Normal,
- Gompertz,
- Gamma models.

Md. Moyazzem Hossain, Prof., ASDS, JU

Exponential Model

In this model, we assume a constant hazard rate over time, and the survival function $S(t)$ can be written as,

$$S(t) = e^{-\lambda t}$$

where λ is the hazard rate.

Md. Moyazzem Hossain, Prof., ASDS, JU

Weibull Model

It allows the hazard rate to change over time and the survival function is,

$$S(t) = e^{-(\lambda t)^{\gamma}}$$

where λ is the scale parameter and γ is the shape parameter.

Log-Normal Model

It is assumed that the logarithm of survival time follows a normal distribution.

The survival function involves the cumulative distribution function of the normal distribution.

Gompertz Model

In this model, the hazard rate increases exponentially with age.

The survival function is expressed in terms of the Gompertz function.

$$f(t) = ae^{-be^{-ct}}$$

Md. Moyazzem Hossain, Prof., ASDS, JU

Gamma Model

It is assumed that survival times follow a Gamma distribution.

The survival function involves the incomplete gamma function.

Md. Moyazzem Hossain, Prof., ASDS, JU

Gamma Model

The upper incomplete gamma function

$$\Gamma(s,x) = \int_x^{\infty} t^{s-1}e^{-t}dt$$

The lower incomplete gamma function

$$\Gamma(s,x) = \int_0^{x} t^{s-1}e^{-t}dt$$

Applications

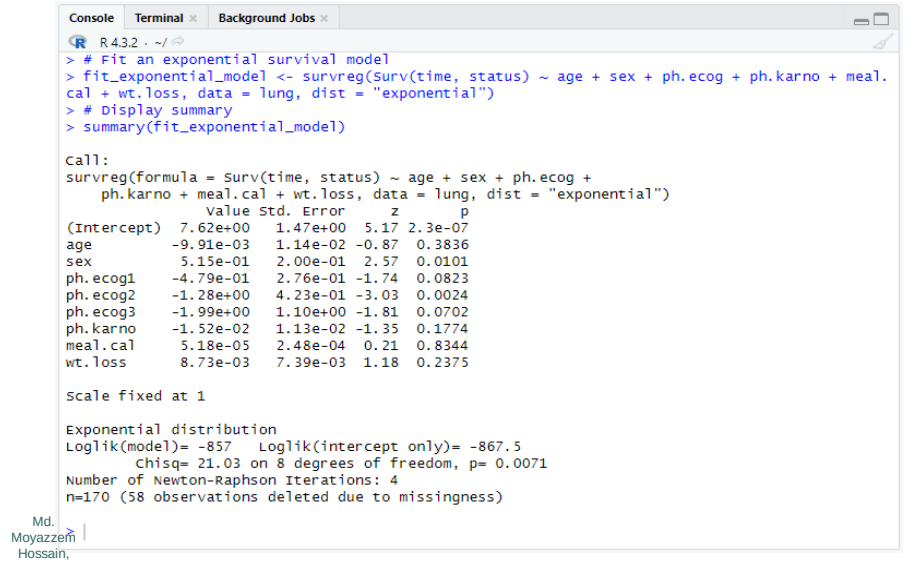
- We consider the "lung" dataset in the `survival` package of R that contains information about survival times and censoring status for patients with advanced lung cancer. In this dataset, the following variables are available:
- inst: Institution code.
- time: This variable represents the survival time or the time until death (measured in days).
- status: This variable indicates the censoring status. A value of 1 represents an observed event (death), and a value of 0 represents censoring (individuals who were still alive at the end of the study).
- sex: The gender of the patient, coded as 1 for male and 2 for female.
- age: The age of the patient at the time of diagnosis.
- ph.ecog: The performance status of the patient, measured on the ECOG scale (Eastern Cooperative Oncology Group). It is a categorical variable representing the overall health and activity level of the patient. Common values include 0 (fully active), 1 (restricted activity but ambulatory), 2 (ambulatory but unable to work), and so on.
- ph.karno: The Karnofsky performance score, another measure of the patient's ability to perform normal daily activities.
- pat.karno: The Karnofsky performance score for the patient's spouse or partner.
- meal.cal: The number of calories consumed during a meal.
- wt.loss: Weight loss in the last six months.

Application_Exponential Model

```
# Fit an exponential survival model
fit_exponential_model <- survreg(Surv(time, status) ~ age + sex + ph.ecog + ph.karno + meal.cal + wt.loss, data = lung, dist = "exponential")

# Display summary
summary(fit_exponential_model)
```

Md. Moyazzem Hossain, Prof., ASDS, JU

Application_Exponential Model



```
> # Fit an exponential survival model
> fit_exponential_model <- survreg(Surv(time, status) ~ age + sex + ph.ecog + ph.karno + meal.cal + wt.loss, data = lung, dist = "exponential")
> # Display summary
> summary(fit_exponential_model)

Call:
survreg(formula = Surv(time, status) ~ age + sex + ph.ecog +
    ph.karno + meal.cal + wt.loss, data = lung, dist = "exponential")
               Value Std. Error     z       p
(Intercept)  7.62e+00   1.47e+00  5.17 2.3e-07
age         -9.91e-03   1.14e-02 -0.87  0.3836
sex          5.15e-01   2.00e-01  2.57  0.0101
ph.ecog1    -4.79e-01   2.76e-01 -1.74  0.0823
ph.ecog2    -1.28e+00   4.23e-01 -3.03  0.0024
ph.ecog3    -1.99e+00   1.10e+00 -1.81  0.0702
ph.karno    -1.52e-02   1.13e-02 -1.35  0.1774
meal.cal     5.18e-05   2.48e-04  0.21  0.8344
wt.loss      8.73e-03   7.39e-03  1.18  0.2375

Scale fixed at 1

Exponential distribution
Loglik(model)= -857   Loglik(intercept only)= -867.5
        Chisq= 21.03 on 8 degrees of freedom, p= 0.0071
Number of Newton-Raphson Iterations: 4
n=170 (58 observations deleted due to missingness)

>
```

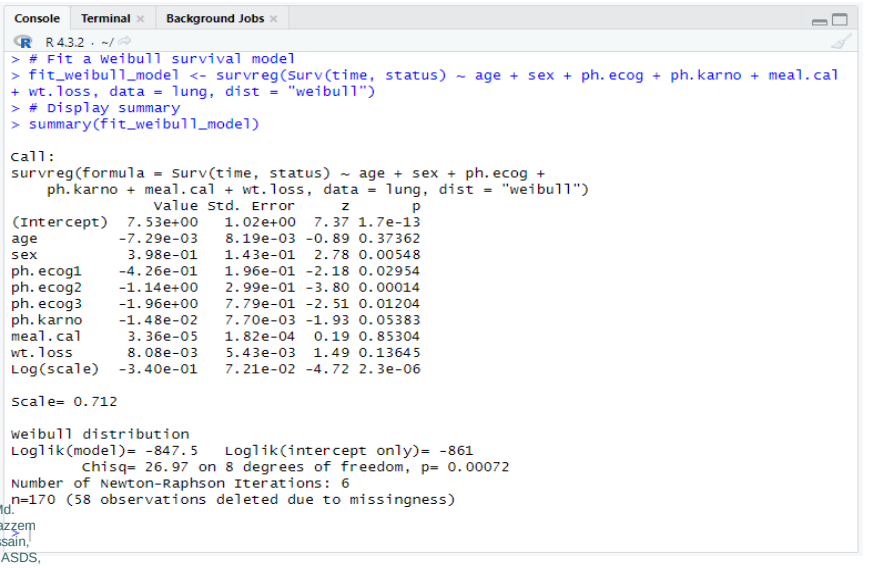
Md. Moyazzem Hossain, Prof., ASDS, JU

Application_Weibull Model

```
# Fit a Weibull survival model
fit_weibull_model <- survreg(Surv(time, status) ~ age + sex + ph.ecog + ph.karno + meal.cal + wt.loss, data = lung, dist = "weibull")

# Display summary
summary(fit_weibull_model)
```

Md. Moyazzem Hossain, Prof., ASDS, JU

Application_Weibull Model



```
> # Fit a Weibull survival model
> fit_weibull_model <- survreg(Surv(time, status) ~ age + sex + ph.ecog + ph.karno + meal.cal
+ wt.loss, data = lung, dist = "weibull")
> # Display summary
> summary(fit_weibull_model)

Call:
survreg(formula = Surv(time, status) ~ age + sex + ph.ecog +
    ph.karno + meal.cal + wt.loss, data = lung, dist = "weibull")
               Value Std. Error     z       p
(Intercept)  7.53e+00   1.02e+00  7.37 1.7e-13
age         -7.29e-03   8.19e-03 -0.89 0.37362
sex          3.98e-01   1.43e-01  2.78 0.00548
ph.ecog1    -4.26e-01   1.96e-01 -2.18 0.02954
ph.ecog2    -1.14e+00   2.99e-01 -3.80 0.00014
ph.ecog3    -1.96e+00   7.79e-01 -2.51 0.01204
ph.karno    -1.48e-02   7.70e-03 -1.93 0.05383
meal.cal     3.36e-05   1.82e-04  0.19 0.85304
wt.loss      8.08e-03   5.43e-03  1.49 0.13645
Log(scale)  -3.40e-01   7.21e-02 -4.72 2.3e-06

Scale= 0.712

Weibull distribution
Loglik(model)= -847.5   Loglik(intercept only)= -861
        Chisq= 26.97 on 8 degrees of freedom, p= 0.00072
Number of Newton-Raphson Iterations: 6
n=170 (58 observations deleted due to missingness)
```

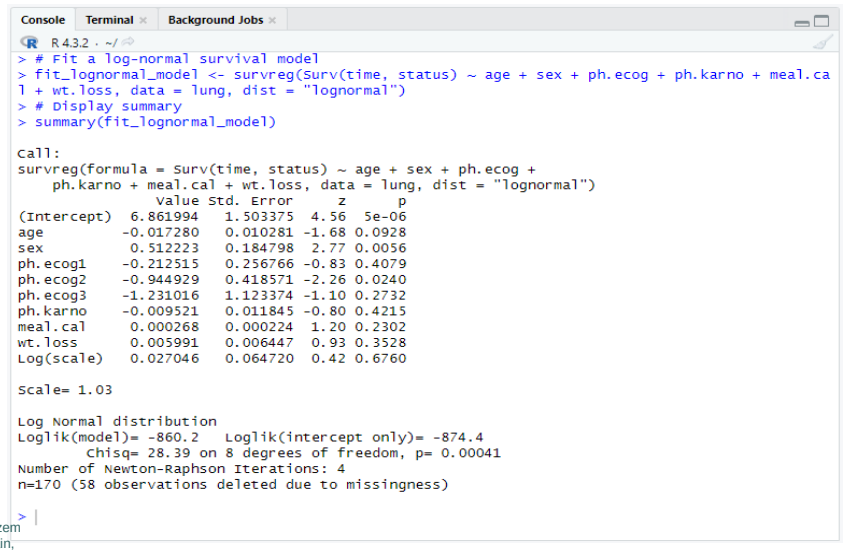
Md. Moyazzem Hossain, Prof., ASDS, JU

Application_Log-normal Model

```
# Fit a log-normal survival model
fit_lognormal_model <- survreg(Surv(time, status) ~ age + sex + ph.ecog + ph.karno + meal.cal + wt.loss, data = lung, dist = "lognormal")

# Display summary
summary(fit_lognormal_model)
```

Md. Moyazzem Hossain, Prof., ASDS, JU

Application_Log-normal Model



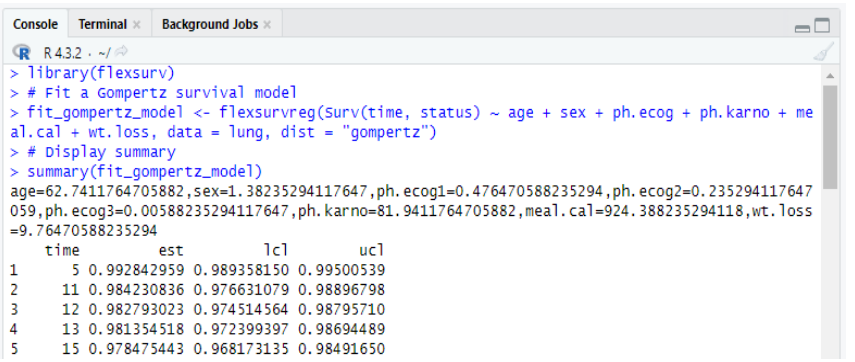
```
Console  Terminal  Background Jobs
R 4.3.2 · ~/
> # Fit a log-normal survival model
> fit_lognormal_model <- survreg(Surv(time, status) ~ age + sex + ph.ecog + ph.karno + meal.cal + wt.loss, data = lung, dist = "lognormal")
> # Display summary
> summary(fit_lognormal_model)

Call:
survreg(formula = Surv(time, status) ~ age + sex + ph.ecog +
    ph.karno + meal.cal + wt.loss, data = lung, dist = "lognormal")
               Value Std. Error     z      p
(Intercept)  6.861994   1.503375  4.56  5e-06
age         -0.017280   0.010281 -1.68 0.0928
sex          0.512223   0.184798  2.77 0.0056
ph.ecog1    -0.212515   0.256766 -0.83 0.4079
ph.ecog2    -0.944929   0.418571 -2.26 0.0240
ph.ecog3    -1.231016   1.123374 -1.10 0.2732
ph.karno    -0.009521   0.011845 -0.80 0.4215
meal.cal     0.000268   0.000224  1.20 0.2302
wt.loss      0.005991   0.006447  0.93 0.3528
Log(scale)   0.027046   0.064720  0.42 0.6760

Scale= 1.03

Log Normal distribution
Loglik(model)= -860.2   Loglik(intercept only)= -874.4
        Chisq= 28.39 on 8 degrees of freedom, p= 0.00041
Number of Newton-Raphson Iterations: 4
n=170 (58 observations deleted due to missingness)

>
```

Md. Moyazzem Hossain, Prof., ASDS, JU

Application_Gompertz Model

```
# Fit a Gompertz survival model
# Install and load the flexsurv package
install.packages("flexsurv")
library(flexsurv)
# Fit a Gompertz survival model
fit_gompertz_model <- flexsurvreg(Surv(time, status) ~ age + sex + ph.ecog + ph.karno + meal.cal + wt.loss, data = lung, dist = "gompertz")
# Display summary
summary(fit_gompertz_model)
fit_gompertz_model$coefficients
fit_gompertz_model$AIC
```


Application_Gompertz Model



Application_Gompertz Model

```
147  814 0.058139625 0.023037392 0.11094251
148  821 0.055355400 0.021398160 0.10692272
149  840 0.048286786 0.017726912 0.09641983
150  965 0.017129008 0.004271564 0.04311755
151 1022 0.009754968 0.001996665 0.02829327
> fit_gompertz_model$coefficients
        shape          rate           age           sex      ph.ecog1      ph.ecog2
 1.944971e-03 -9.187697e+00  1.161597e-02 -5.413610e-01  6.419393e-01  1.690979e+00
     ph.ecog3      ph.karno      meal.cal       wt.loss
 2.752043e+00  2.512697e-02 -2.254047e-05 -1.035909e-02
> fit_gompertz_model$AIC
[1] 1715.584
>
```

Md. Moyazzem Hossain, Prof., ASDS, JU

Application_Gamma Model

```
# Fit a gamma survival model
fit_gamma_model <- flexsurvreg(Surv(time, status) ~ age + sex + ph.ecog + ph.karno + wt.loss, data = lung, dist = "gamma")

# Display summary
summary(fit_gamma_model)
fit_gamma_model$coefficients
fit_gamma_model$AIC
```

Md. Moyazzem Hossain, Prof., ASDS, JU

Application_Gamma Model

```
Console  Terminal  Background Jobs
R 4.3.2 · ~/
> # Fit a gamma survival model
> fit_gamma_model <- flexsurvreg(Surv(time, status) ~ age + sex + ph.ecog + ph.karno  + wt.l
oss, data = lung, dist = "gamma")
> # Display summary
> summary(fit_gamma_model)
age=62.5023474178404,sex=1.4037558685446,ph.ecog1=0.497652582159624,ph.ecog2=0.2112676056338
03,ph.ecog3=0.00469483568075117,ph.karno=82.1596244131455,wt.loss=9.72769953051643
    time        est        lcl        ucl
1      5 0.99897976 0.99705031 0.99971859
2     11 0.99627645 0.99139772 0.99863102
3     12 0.99570823 0.99032545 0.99837171
4     13 0.99511024 0.98922306 0.99809052
5     15 0.99382925 0.98691159 0.99746263
```

Md. Moyazzem Hossain, Prof., ASDS, JU

Application_Gamma Model

```
174  840 0.09664191 0.07208869 0.12031129
175  883 0.08327248 0.06048074 0.10637355
176  965 0.06251939 0.04292331 0.08389162
177 1010 0.05334659 0.03546415 0.07368652
178 1022 0.05112882 0.03369212 0.07117627
> fit_gamma_model$coefficients
       shape         rate          age          sex     ph.ecog1     ph.ecog2
 0.507571209 -6.537342250  0.009890963 -0.465749975  0.428972854  1.000110739
    ph.ecog3     ph.karno      wt.loss
 1.664281226  0.008481681 -0.006275392
> fit_gamma_model$AIC
[1] 2112.201
>
```

Md. Moyazzem Hossain, Prof., ASDS, JU

Thank You

Md.
Moyazzem
Hossain,
Prof., ASDS,
JU