KM

parametric survival function (for 6-mp group)

```
data(gehan)
Surv(gehan$time,gehan$cens,type='right') # 0: censored, 1: observed
```

```
## [1]  1  10  22   7   3  32+ 12  23   8  22  17   6   2  16  11  34+  8  32+ 12
## [20] 25+  2  11+  5  20+  4  19+ 15   6   8  17+ 23  35+  5   6  11  13   4   9+
## [39]  1   6+  8  10+
```

```
param1 <- flexsurvreg(Surv(time, cens) ~ 1, data = subset(gehan, treat=="6-MP"),
                      dist = "exp")   # S(t)=e^{-rate*t}

param2 <- flexsurvreg(Surv(time, cens) ~ 1, data = subset(gehan, treat=="6-MP"),
                      dist = "weibull") # S(t)=e^{-(t/scale)^shape}

param2 # Weibull parameter estimation and CI
```

```
## Call:
## flexsurvreg(formula = Surv(time, cens) ~ 1, data = subset(gehan,
##     treat == "6-MP"), dist = "weibull")
##
## Estimates:
##         est      L95%     U95%     se
## shape   1.354    0.784    2.336    0.377
## scale   33.765   19.760   57.698   9.230
##
## N = 21,  Events: 9,  Censored: 12
## Total time at risk: 359
## Log-likelihood = -41.65868, df = 2
## AIC = 87.31736
```

```
summary(param2) # survival function estimation and CI
```

```
##
##     time       est         lcl         ucl
## 1      6 0.9080619 0.78773350 0.9799141
## 2      7 0.8879657 0.75822302 0.9707331
## 3      9 0.8462212 0.69808986 0.9474850
## 4     10 0.8248347 0.66641800 0.9358391
## 5     11 0.8032471 0.63087380 0.9255711
## 6     13 0.7598077 0.56872314 0.8966506
## 7     16 0.6949980 0.46108564 0.8441076
```

```
## 8      17 0.6737026 0.42641851 0.8256091
## 9      19 0.6318222 0.36799388 0.7939847
## 10     20 0.6113031 0.33500947 0.7794978
## 11     22 0.5712422 0.26874916 0.7467204
## 12     23 0.5517438 0.24361526 0.7317714
## 13     25 0.5138977 0.19017378 0.7008651
## 14     32 0.3945962 0.06595602 0.5980620
## 15     34 0.3644276 0.04462528 0.5773773
## 16     35 0.3499986 0.03631243 0.5673157
```
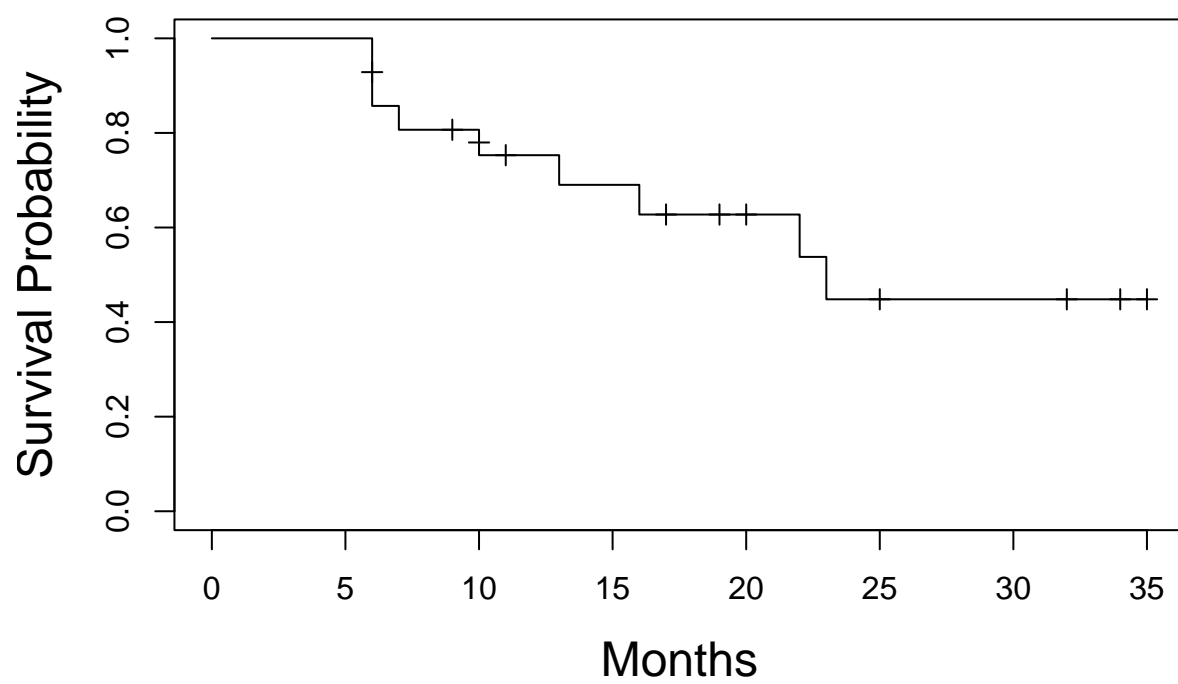
```
plot(param2, xlab="Months", ylab="Survival Probability", main="6MP (KM and Parametric Est with 95% CI)"
```
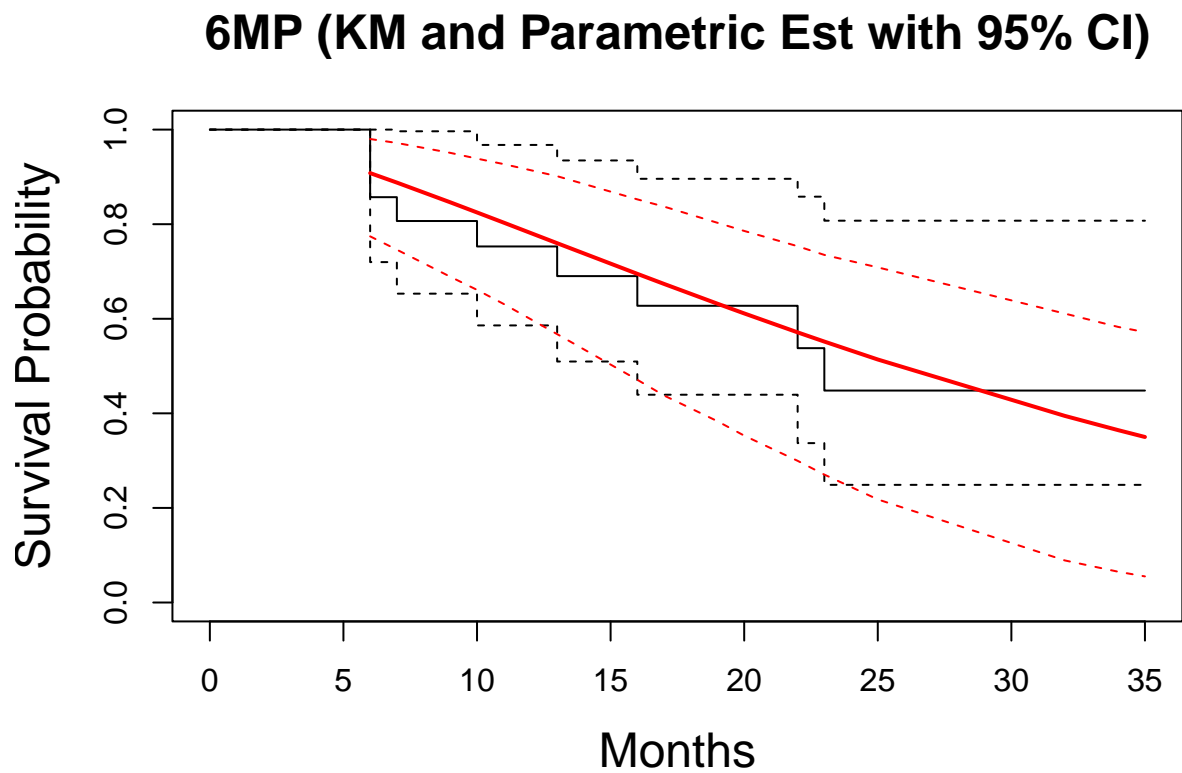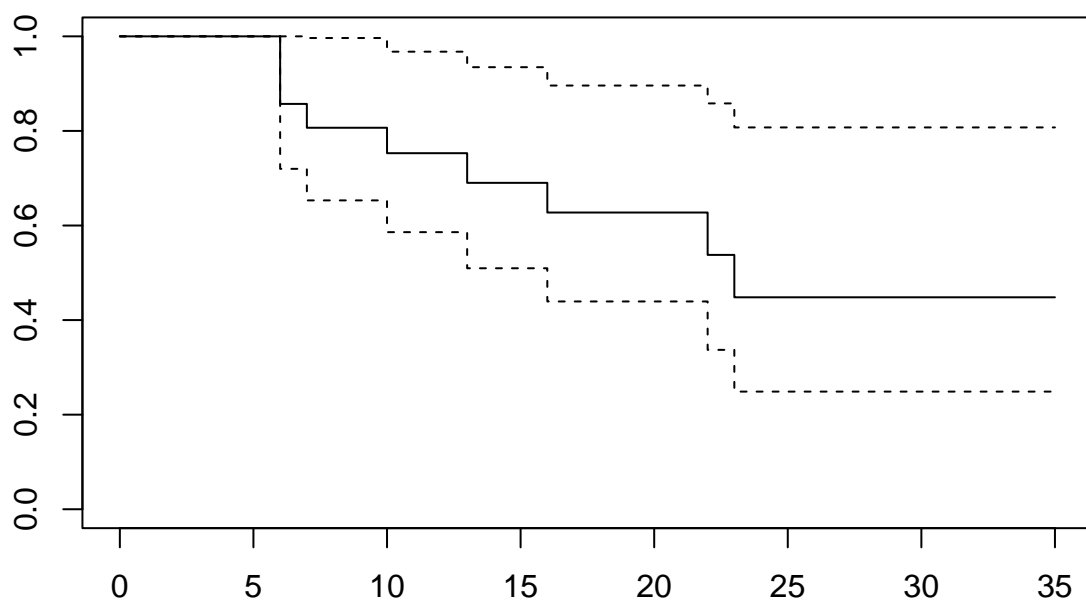

KM survival function

```
KM = survfit(Surv(time,cens)~1, data = subset(gehan, treat=="6-MP"), conf.type='log')
plot(KM, conf.int = FALSE, mark.time = TRUE,xlab="Months",
     ylab="Survival Probability", main="6MP K-M curve", cex.lab = 1.5, cex.main = 1.5)
```
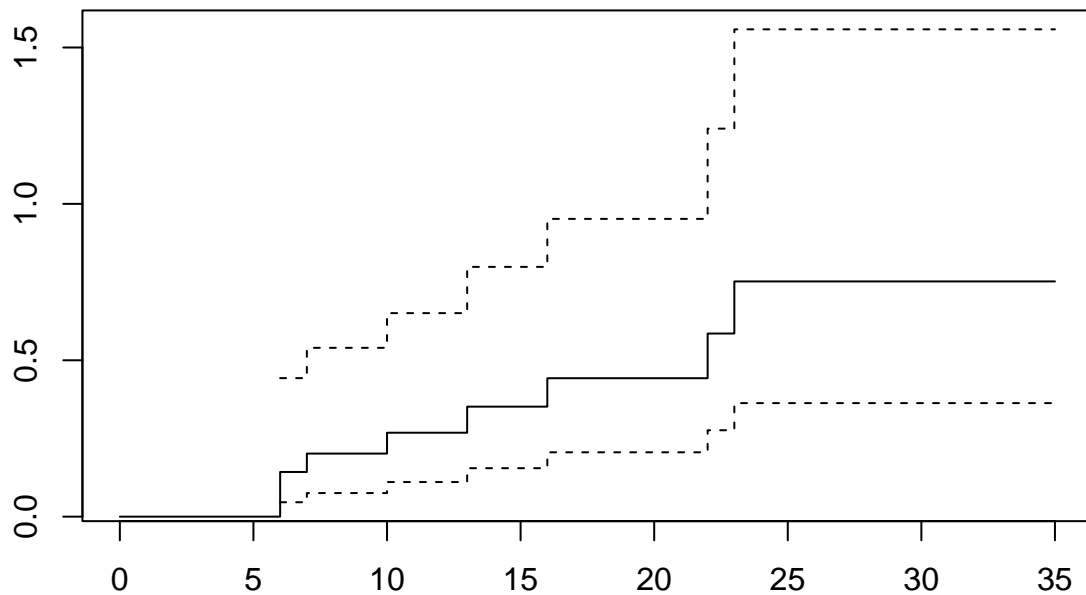
6MP K–M curve


```
plot(KM)
```

```
plot(KM,fun='cumhaz') # cumulative hazard fun
```

```
# estimate cumulative hazard rates
cbind(KM$time,-log(KM$surv),  cumsum(KM$n.event/KM$n.risk)) # time, KM est, Nelson-Aalen Estimator
```

```
##        [,1]      [,2]      [,3]
## [1,]     6 0.1541507 0.1428571
## [2,]     7 0.2147753 0.2016807
## [3,]     9 0.2147753 0.2016807
## [4,]    10 0.2837682 0.2683473
## [5,]    11 0.2837682 0.2683473
## [6,]    13 0.3707796 0.3516807
## [7,]    16 0.4660897 0.4425898
## [8,]    17 0.4660897 0.4425898
## [9,]    19 0.4660897 0.4425898
## [10,]    20 0.4660897 0.4425898
## [11,]    22 0.6202404 0.5854469
## [12,]    23 0.8025620 0.7521136
## [13,]    25 0.8025620 0.7521136
## [14,]    32 0.8025620 0.7521136
## [15,]    34 0.8025620 0.7521136
## [16,]    35 0.8025620 0.7521136
```

```
#
# obtain survival rate at given time, with CI
summary(KM,time=c(5,10,12.5, 15)) # note: n.event is the cumulative num of events since last listed tim
```

```
## Call: survfit(formula = Surv(time, cens) ~ 1, data = subset(gehan,
```

```
##      treat == "6-MP"), conf.type = "log")
##
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##   5.0     21       0    1.000  0.0000        1.000        1.000
## 10.0     15       5    0.753  0.0963        0.586        0.968
## 12.5     12       0    0.753  0.0963        0.586        0.968
## 15.0     11       1    0.690  0.1068        0.510        0.935
```

```r
summary(KM, censored = TRUE)# (if not specify time, then n.event is the # event at each time point)
```

```
## Call: survfit(formula = Surv(time, cens) ~ 1, data = subset(gehan,
##      treat == "6-MP"), conf.type = "log")
##
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##     6     21       3    0.857  0.0764        0.720        1.000
##     7     17       1    0.807  0.0869        0.653        0.996
##     9     16       0    0.807  0.0869        0.653        0.996
##    10     15       1    0.753  0.0963        0.586        0.968
##    11     13       0    0.753  0.0963        0.586        0.968
##    13     12       1    0.690  0.1068        0.510        0.935
##    16     11       1    0.627  0.1141        0.439        0.896
##    17     10       0    0.627  0.1141        0.439        0.896
##    19      9       0    0.627  0.1141        0.439        0.896
##    20      8       0    0.627  0.1141        0.439        0.896
##    22      7       1    0.538  0.1282        0.337        0.858
##    23      6       1    0.448  0.1346        0.249        0.807
##    25      5       0    0.448  0.1346        0.249        0.807
##    32      4       0    0.448  0.1346        0.249        0.807
##    34      2       0    0.448  0.1346        0.249        0.807
##    35      1       0    0.448  0.1346        0.249        0.807
```

```r
# median survival time, with CI
print(KM)
```

```
## Call: survfit(formula = Surv(time, cens) ~ 1, data = subset(gehan,
##      treat == "6-MP"), conf.type = "log")
##
##        n  events  median 0.95LCL 0.95UCL
##       21       9      23      16      NA
```

Log Rank test

```r
survdiff(Surv(time,cens)~treat, data=gehan) # log rank test
```

```
## Call:
## survdiff(formula = Surv(time, cens) ~ treat, data = gehan)
##
##                 N Observed Expected (O-E)^2/E (O-E)^2/V
## treat=6-MP     21        9     19.3      5.46      16.8
## treat=control 21       21     10.7      9.77      16.8
##
##  Chisq= 16.8  on 1 degrees of freedom, p= 4e-05
```

```
plot(survfit(Surv(time,cens)~treat, data = gehan))
library(survminer)
```

```
## Loading required package: ggplot2
```

```
## Loading required package: ggpubr
```

```
ggsurvplot( survfit(Surv(time, cens) ~ treat, data = gehan), conf.int=TRUE)
```