Survival Analysis of Chronic Active Hepatitis Patients with Prednisolone Treatment

Nick Wibert

Preface

In a clinical trial described by Kirk et al. (1980), 44 patients with chronic active hepatitis were randomized to the drug *prednisolone*, or to an untreated control group. The survival time of the patients, in months, following admission to the trial, was the response variable of interest.

The Random Right Censorship Model

In a survival study, right-censored data refers to the individuals whose time-to-event (usually death) was not observed for any reason. The mathematical model to account for right-censored data, known as the *random right censorship model*, includes two random variables,

$$X_i, ..., X_n \overset{\text{iid}}{\sim} F$$
$$Y_i, ..., Y_n \overset{\text{iid}}{\sim} G$$

where X_i represents time-to-event (survival) times and Y_i represents censoring times. The model also defines two additional variables using X_i and Y_i:

$$Z_i = \min\{X_i, Y_i\}$$
$$\delta = I(X_i \leq Y_i)$$

where Z_i is simply the observed time (either survival or censoring time, whichever occurred first) and δ_i is an indicator variable which takes on the value 1 when time-to-event is observed and 0 otherwise.

Let's look at a brief example to demonstrate what this notation means. In the Kirk study, two of the observed times for the prednisolone group were 2 and 56^+ ($^+$ denotes a censored observation). Suppose these are our only two observations. For the first observation ($i = 1$), the time-to-event was observed, so we assign $X_1 = 2$ and $Y_1 = (2, \infty)$ since no censoring time is observed. The second observation ($i = 2$) has an observed censoring time, so $X_2 = (52, \infty)$ and $Y_2 = 52$. Furthermore we have

$$Z_1 = \min\{X_1, Y_1\} = X_1 = 2$$
$$Z_2 = \min\{X_2, Y_2\} = Y_2 = 52$$
$$\delta_1 = I(X_1 \leq Y_1) = 1$$
$$\delta_2 = I(X_2 \leq Y_2) = 0$$

Z_i always takes on the value that was observed, and δ_i simply tells us the "status" of data point—whether or not it is censored. Thus, everything we need to know about an individual observation is described by these two variables alone.

An important assumption of the random right censorship model is that the variables X_i and Y_i are independent. Here are some common situations in survival studies that involve this fact:

- Subject dies of a cause unrelated to the condition being studied
  - In this scenario, the independence assumption is met; survival time is observed and not influenced by censoring time.
- Subject withdraws from the study due to side effects from the treatment
  - Here the independence assumption is violated. Censoring time and survival time are simultaneously influenced by treatment in this case.
- Subject moves away and cannot continue in the study
  - The assumption is met, as the censoring time is influenced by factors unrelated to the study.

The Kaplan-Meier Estimator (KME)

When analyzing ungrouped censored survival data, the first step is usually to compute the *Kaplan-Meier estimate (KME)* of the survivor function $S(t)$. This computation involves constructing a series of time intervals for the data such that one death occurs in each interval, taken to occur at the start of the interval.

Using the product-limit approach, one moves through the observed times in ascending order, evenly redistributing the "weight" of each censored time (proportion of interval) to each observed time-to-event (death). Once the end of the dataset is reached, an empirical estimate of the survivor function (probability an individual survives to the end of the study) is obtained. This can be computed in a tabular form using the variables t_i, n_i, d_i, and c_i. Let n_i denote the number of "at-risk" individuals entering time t_i, while d_i and c_i denote the number of deaths and number of censored observations which occurred at time t_i.

We can also obtain the variance of the KME at each observation using *Greenwood's formula*, written as

$$\hat{\sigma}^2_G(t) = (\hat{S}(t))^2 \cdot \sum_{i:Z_{(i)} \leq t} \delta_{(i)} \frac{d_i}{(n_i - d_i)n_i}$$

which is made up only of variables we have already defined.

Let's do another brief example using a subset of observations from the prednisolone group in the Kirk study. Suppose that the five observed times {12, 54, 56^+, 68, 89} are the entire dataset. The Kaplan-Meier estimate and Greenwood estimates for variance at each level are calculated in the table below.

t_i	n_i	d_i	c_i	$(n_i - d_i)/n_i$	$\hat{S}(t_i)$	$\frac{d_i}{(n_i - d_i)n_i}$	$\hat{\sigma}^2_G(t_i)$
0	5	0	0	1.00	1.00	-	-
12	5	1	0	0.80	0.80	0.050	0.032
54	4	1	0	0.75	0.60	0.083	0.048
56^+	3	0	1	1.00	0.60	0.000	0.048
68	2	1	0	0.50	0.30	0.500	0.057
89	1	1	0	0.00	0.00	-	-

Since the final interval contains only one individual left alive, and their death is observed, the final estimate of the survivor function is 0 (all individuals either died by the final interval, or were censored before reaching the final interval). Also note that since the final observation is uncensored, the Greenwood estimate of variance is undefined at that point.

Estimating the Survivor Function for the Prednisolone Study

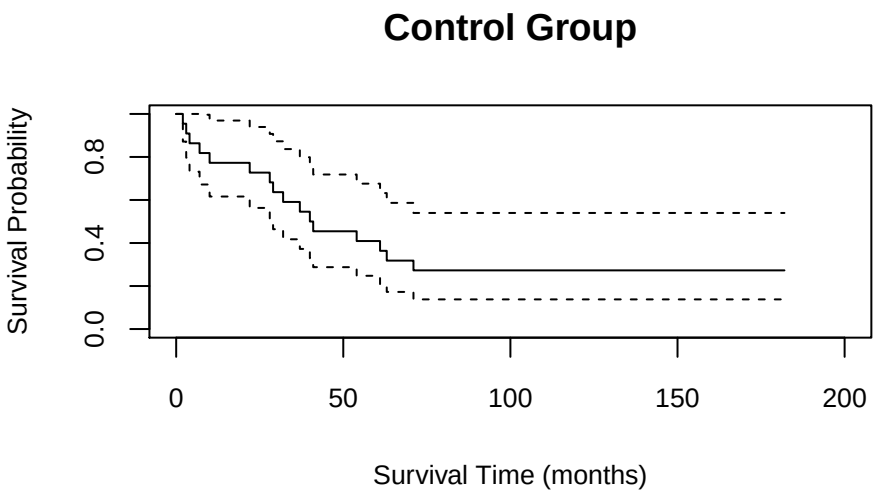
Now that we have clearly laid out the random right censorship model and the Kaplan-Meier estimator, we are equipped to compute an estimate for the survivor function of the prednisolone study by Kirk et al. (1980).

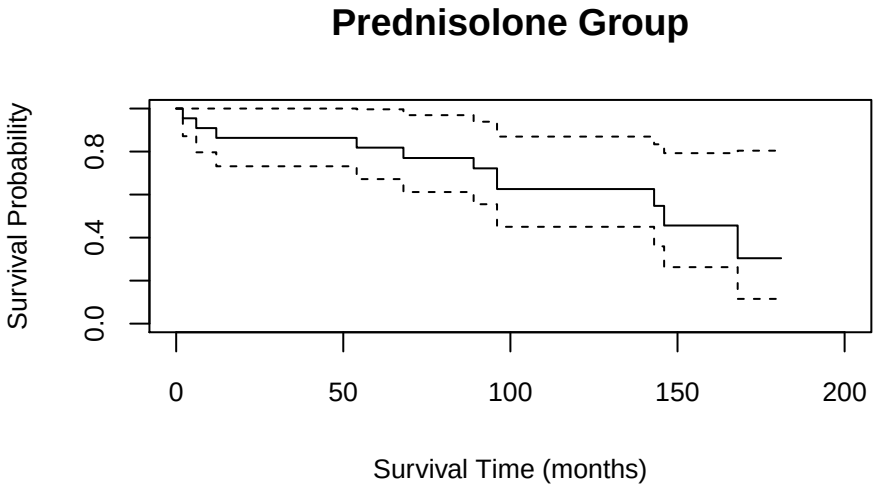
As aforementioned, the 44 chronic active hepatitis patients who participated in the study were randomly assigned to two groups: a prednisolone treatment group, or an untreated control group. The response variable of interest is the survival time following admission to the trial.

Using the `survival` library in R, we can obtain the KME for each of the two groups separately. In the code below, the dataset is read in, and the survivor functions are estimated using the `survfit` function.

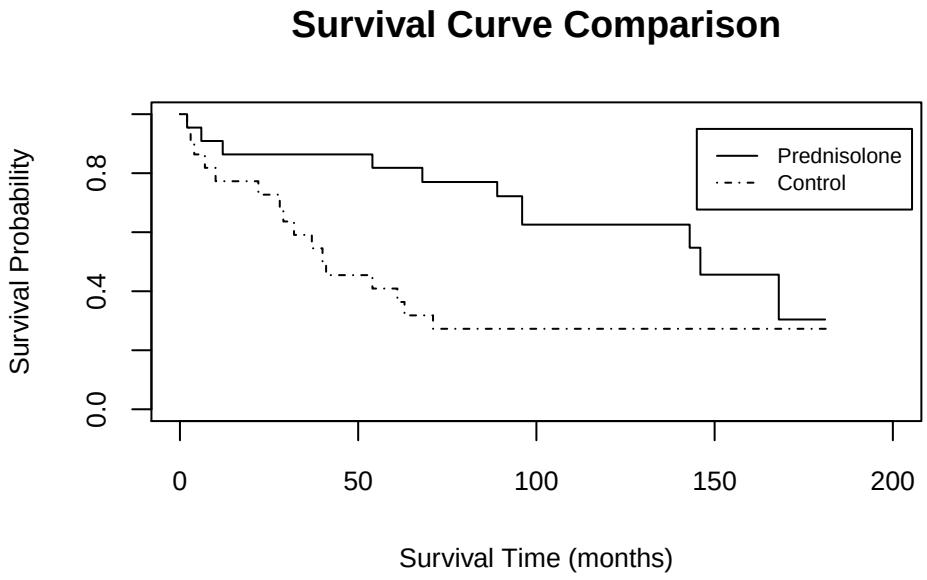
```
hep <- read.table("Chronic-active-hepatitis.dat", header=TRUE)
# treatment 1 = prednisolone, treatment 2 = control
surv.ctrl <- survfit(Surv(time,status) ~ 1, data=hep, subset=treatment==2)
surv.trt <- survfit(Surv(time,status) ~ 1, data=hep, subset=treatment==1)
```

These two estimates for the survivor function are plotted separately below, with dotted lines representing the 95% confidence bands for the estimate.





The two survival curves illustrate noticeable differences in the trend of the control group vs. the treatment group. For an easier comparison, let's look at a plot with the two curves on the same graph, and remove the confidence bands for less visual clutter.



The probability of survival in the control group drops much earlier (hitting its lowest point at around 70 months) than in the prednisolone group, where we see a steadier decline. This suggests that prednisolone therapy may improve the odds of survival for a chronic active hepatitis patient. Survival probability is very similar by the end of the trial in both groups (~0.3), suggesting that prednisolone treatment may become less effective over time at increasing your odds of survival compared to no treatment.

It is worth noting that the last observed death for the prednisolone group was at 168 months, while the last observed death for the control group was at only 71 months (all following observations are censored). The prednisolone group also has almost twice as many censored individuals as the control (11 vs. 6), which could possibly exaggerate the observed effects of the drug on survival time.

The median survival time for each group may also be of interest, which we can obtain with the `print` function:

```
print(surv.ctrl); print(surv.trt)
```

```
## Call: survfit(formula = Surv(time, status) ~ 1, data = hep, subset = treatment ==
##     2)
##
##       n  events  median 0.95LCL 0.95UCL
##    22.0    16.0    40.5    29.0      NA
```

```
## Call: survfit(formula = Surv(time, status) ~ 1, data = hep, subset = treatment ==
##     1)
##
##       n  events  median 0.95LCL 0.95UCL
##      22      11     146      96      NA
```

Observe that the median survival time is 40.5 months for the control group and 146 months for the prednisolone group, more possible evidence for improved survival time when prescribed with the drug.

Let's compare the estimates further with more specific, numerical analysis. For example, what is the probability of surviving past the 36-month mark for each group ($S(36)$)? Using the `summary` function in R, we can determine the value of the survivor function at month 36 for each group.

```
summary(surv.ctrl); summary(surv.trt)
```

```
## Call: survfit(formula = Surv(time, status) ~ 1, data = hep, subset = treatment ==
##     2)
##
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##     2     22       1    0.955  0.0444        0.871        1.000
##     3     21       1    0.909  0.0613        0.797        1.000
##     4     20       1    0.864  0.0732        0.732        1.000
##     7     19       1    0.818  0.0822        0.672        0.996
##    10     18       1    0.773  0.0893        0.616        0.969
##    22     17       1    0.727  0.0950        0.563        0.939
##    28     16       1    0.682  0.0993        0.513        0.907
##    29     15       1    0.636  0.1026        0.464        0.873
##    32     14       1    0.591  0.1048        0.417        0.837
##    37     13       1    0.545  0.1062        0.372        0.799
##    40     12       1    0.500  0.1066        0.329        0.759
##    41     11       1    0.455  0.1062        0.288        0.718
##    54     10       1    0.409  0.1048        0.248        0.676
##    61      9       1    0.364  0.1026        0.209        0.632
##    63      8       1    0.318  0.0993        0.173        0.587
##    71      7       1    0.273  0.0950        0.138        0.540
```

```
## Call: survfit(formula = Surv(time, status) ~ 1, data = hep, subset = treatment ==
##     1)
##
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##     2     22       1    0.955  0.0444        0.871        1.000
##     6     21       1    0.909  0.0613        0.797        1.000
##    12     20       1    0.864  0.0732        0.732        1.000
##    54     19       1    0.818  0.0822        0.672        0.996
##    68     17       1    0.770  0.0904        0.612        0.969
##    89     16       1    0.722  0.0967        0.555        0.939
##    96     15       2    0.626  0.1051        0.450        0.870
##   143      8       1    0.547  0.1175        0.359        0.834
##   146      6       1    0.456  0.1285        0.263        0.793
##   168      3       1    0.304  0.1509        0.115        0.804
```

From the output above, we see that the estimates are given by

$$S_{ctrl}(36) = P(X_{ctrl} \geq 36) = 0.591$$
$$S_{trt}(36) = p(X_{trt} \geq 36) = 0.864$$

The output also provides 95% confidence intervals for each survival estimate:

$$S_{ctrl}(36) \in [0.417, 0.837]$$
$$S_{trt}(36) \in [0.732, 1.000]$$

Therefore we are 95% confident that the true survival probability for the control group lies on the interval $[0.417, 0.837]$ and the true survival probability for the prednisolone group lies on the interval $[0.732, 1]$.

Based on these estimates, the probability of surviving past 36 months for a patient treated with prednisolone appears to be significantly higher than for an untreated patient. We should note, however, that the confidence intervals overlap, and thus we cannot draw any definitive conclusions unless we perform a test.

Confidence intervals for median survival time (Exponential model)

Let's fit an Exponential model to each of the two groups separately. From these models we can calculate an estimate and 95% confidence interval for the median survival time within each group.

```
hep <- read.table("Chronic-active-hepatitis.dat", header=TRUE)
hep$treatment <- factor(hep$treatment, levels=1:2, labels=c("Pred","Ctrl"))

exp.pred <- survreg(Surv(time,status) ~ 1, data=hep, subset=treatment=="Pred",
                    dist="exponential")
exp.ctrl <- survreg(Surv(time,status) ~ 1, data=hep, subset=treatment=="Ctrl",
                    dist="exponential")
```

First, we must compute the MLE of λ using the formula

$$\hat{\lambda} = \frac{n_u}{\sum_{i=1}^{n} Z_i}$$

where $n_u = \sum_{i=1}^{n} \delta_i$ is the total number of uncensored observations.

Constructing median confidence intervals in R

```
# split data based on treatment; can now access groups using hep2$Pred and hep2$Ctrl
hep2 <- split(hep, hep$treatment)
```

First, we'll construct an interval for the control group:

```
## MEDIAN INTERVAL FOR CONTROL GROUP
# calculate MLE of lambda
n_u <- sum(hep2$Ctrl$status)
Z <- sum(hep2$Ctrl$time)
MLE <- n_u / Z; MLE
```

```
## [1] 0.01123596
```

```
# estimate median survival time
med <- log(2) / MLE; med
```

```
## [1] 61.6901
```

```
# log standard error of median estimate
SE <- 1 / sqrt(n_u); SE
```

```
## [1] 0.25
```

```
# create log confidence interval, and exponentiate to get true interval for median
z <- qnorm(1-(0.05/2))
lower <- log(med) - z*SE; upper <- log(med) + z*SE
ctrl.interval <- c(exp(lower), exp(upper))
```

Next, we'll do the same for the treatment group:

```
## MEDIAN INTERVAL FOR TREATMENT GROUP
# calculate MLE of lambda
n_u <- sum(hep2$Pred$status)
Z <- sum(hep2$Pred$time)
MLE <- n_u / Z; MLE
```

```
## [1] 0.004564315
```

```
# estimate median survival time
med <- log(2) / MLE; med
```

```
## [1] 151.8622
```

```
# log standard error of median estimate
SE <- 1 / sqrt(n_u); SE
```

```
## [1] 0.3015113
```

```
# create log confidence interval, and exponentiate to get true interval for median
z <- qnorm(1-(0.05/2))
lower <- log(med) - z*SE; upper <- log(med) + z*SE
pred.interval <- c(exp(lower), exp(upper))

# print final intervals
ctrl.interval; pred.interval
```

```
## [1]  37.79332 100.69684
```

```
## [1]  84.10134 274.21849
```

We are 95% confident that the true median survival time lies on the interval $[37.8, 100.7]$ for the *control* group, and $[84.1, 274.2]$ for the *prednisolone* group.

Mantel-Haenszel test with a subset of the data

We wish to compare the survival distributions for the two groups in the study. A good way to do this is to perform the Mantel-Haenszel test for equality.

We will perform the Mantel-Haenszel test of equality in a chart format; but first, let's define a few variables that will be useful in the calculations. The deaths and risk sets at point t can be represented as a table:

	No. Dead	No. Alive
Prednisolone	D_1	$R_1 - D_1$
Control	D_2	$R_2 - D_2$

where D_i is the number of deaths in population i at time t, and R_i is the number alive in population i at the start of time t (the risk set). The null and alternative hypotheses are

$$H_0 : A_1(t) = A_2(t)$$
$$H_a : A_1(t) \neq A_2(t)$$

where $A_i(t)$ is the cumulative hazard function for population i.

Our test is concerned with D_1, which has a hypergeometric distribution if the null hypothesis is true. The mean and variance of this D_1 under the null hypothesis are defined as

$$E_{0,M}(D_1(t)) = (D_1 + D_2)\frac{R_1}{R_1 + R_2}$$
$$\text{Var}_{0,M}(D_1(t)) = (D_1 + D_2)\frac{R_1}{R_1 + R_2}\frac{R_2}{R_1 + R_2} = E_{0,M}(D_1(t))\frac{R_2}{R_1 + R_2}$$

and finally, we use these values to compute the Mantel-Haenszel test statistic,

$$MH = \frac{\sum_t [D_1(t) - E_{0,M}(D_1(t))]}{\sqrt{\sum_t \text{Var}_{0,M}(D_1(t))}}$$

though the statistic we will ultimately report is MH^2 due to the fact that it follows a chi-squared distribution under the null. All of the calculations are shown in the table below.

Time (t)	$D_1(t)$	$D_2(t)$	$R_1(t)$	$R_2(t)$	$E_{0,M}(D_1(t))$	$\text{Var}_{0,M}(D_1(t))$	$D_1(t) - E_{0,M}(D_1(t))$
54	1	0	5	5	$(1+0)\left(\frac{5}{5+5}\right) = \frac{1}{2}$	$\left(\frac{1}{2}\right)\left(\frac{5}{5+5}\right) = \frac{1}{4}$	$\frac{1}{2}$
55	0	1	4	5	$(0+1)\left(\frac{4}{4+5}\right) = \frac{4}{9}$	$\left(\frac{4}{9}\right)\left(\frac{5}{4+5}\right) = \frac{20}{81}$	$-\frac{4}{9}$
56	0	0	4	4	0	0	0
61	0	1	3	4	3/7	12/49	$-3/7$
63	0	1	3	3	1/2	1/4	$-1/2$
68	1	0	3	2	3/5	6/25	2/5
71	0	1	2	2	1/2	1/4	$-1/2$
125	0	0	2	1	0	0	0
140	0	0	1	1	0	0	0
143	1	0	1	0	1	0	0

$$MH^2 = \frac{(\sum_t [D_1(t) - E_{0,M}(D_1(t))])^2}{\sum_t \text{Var}_{0,M}(D_1(t))} = \frac{(613/630)^2}{588131/396900} \approx \boxed{0.639}$$

The null distribution of MH^2 is a chi-squared distribution with 1 df. Therefore the p-value of our test is $P(\chi_1^2 > 0.639) \approx 0.42$ and we fail to reject the null hypothesis. There is *not* sufficient evidence to conclude that the survival distributions of the two groups are different.

Gehan test with a subset of the data

Form of Gehan test statistic is given by

$$U = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left[I\left(Z_j^{(2)} \geq Z_i^{(1)}\right) \delta_i^{(1)} - I\left(Z_j^{(2)} \leq Z_i^{(1)}\right) \delta_i^{(2)} \right]$$

where $Z_i^{(1)}$ denotes the observations from the first population and $Z_i^{(2)}$ denotes observations from the second.

Let the prednisolone group be population 1, and the control group be population 2 (as we did in the Mantel-Haenszel test).

```
U <- 0  # initialize Gehan statistic
pred <- data.frame("time" = c(54,56,68,125,143),
                   "status" = c(1,0,1,0,1))

ctrl <- data.frame("time" = c(55,61,63,71,140),
                   "status" = c(1,1,1,1,0))

for (i in 1:5)
{
  for (j in 1:5)
  {
    # I1 and I2 represent the two indicator terms
    I1 <- ifelse(ctrl$time[j] >= pred$time[i], 1, 0)
    I2 <- ifelse(ctrl$time[j] <= pred$time[i], 1, 0)

    U <- U + (I1*pred$status[i] - I2*ctrl$status[j])
  }
}

U   # print Gehan statistic
```

```
## [1] -5
```

Two Procedures for Comparing Survival Distributions

In this section, we will carry out two hypothesis test procedures to compare the survival distributions for the Prednisolone group and the Control group (using the entire dataset this time). We will implement the `survdiff()` function from the R "survival" package to perform the tests.

We use the same hypotheses from earlier for both tests:

$$H_0 : A_1(t) = A_2(t)$$
$$H_a : A_1(t) \neq A_2(t)$$

where $A_i(t)$ is the cumulative hazard function for population i.

```
# log-rank (Mantel-Haenszel test)
survdiff(Surv(time,status) ~ treatment, data=hep, rho=0)
```

```
## Call:
## survdiff(formula = Surv(time, status) ~ treatment, data = hep,
##     rho = 0)
##
##                 N Observed Expected (O-E)^2/E (O-E)^2/V
## treatment=Pred 22       11     16.4      1.77      4.66
## treatment=Ctrl 22       16     10.6      2.73      4.66
##
##  Chisq= 4.7  on 1 degrees of freedom, p= 0.03
```

```
# Prentice/Peto-Peto modification of the Gehan test
survdiff(Surv(time,status) ~ treatment, data=hep, rho=1)
```

```
## Call:
## survdiff(formula = Surv(time, status) ~ treatment, data = hep,
##     rho = 1)
##
##                 N Observed Expected (O-E)^2/E (O-E)^2/V
## treatment=Pred 22     6.72    11.17      1.78      5.85
## treatment=Ctrl 22    12.29     7.84      2.53      5.85
##
##  Chisq= 5.8  on 1 degrees of freedom, p= 0.02
```

Observe that the MH test statistic is 4.7 while the P/P-P modification of the Gehan test statistic is 5.8. Under our standard confidence level of $\alpha = 0.05$, both test results would lead us to reject the null hypothesis, and conclude that prednisolone treatment has an effect on survival time. In other words, there is sufficient evidence to conclude that the survival distributions for the two groups (Prednisolone and Control) are different.

While both tests provided similar results in this case, we should use our "rough rule-of-thumb" for which test is better in this scenario. This rule of thumb states that if the survival curves *don't* cross, use the Mantel-Haenszel test; if they *do* cross, use the Gehan test.

Recall from our survival curve plot on page 4 that the survival curves for these two groups *don't* cross; therefore, we would probably want to perform the MH test over the Gehan in this setting.