Survival _II

Kelvin Nyongesa

2023-06-08

```
# Load the survival package
library(survival)

# Filter the dataset to include only observed events
kidney_observed <- kidney[kidney$status == 1, ]

# Create a survival object with the observed times
surv_obj <- Surv(kidney_observed$time)

# Fit the Kaplan-Meier estimator
km_fit <- survfit(surv_obj ~ 1)

# Access the estimated survival probabilities at each observed time point
km_survival <- km_fit$surv

# Print the estimated survival probabilities
library(knitr)
km_table<-summary(km_fit)
km_table
```

```
## Call: survfit(formula = surv_obj ~ 1)
##
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##     2     58       1   0.9828  0.0171      0.94982        1.000
##     7     57       2   0.9483  0.0291      0.89296        1.000
##     8     55       2   0.9138  0.0369      0.84434        0.989
##     9     53       1   0.8966  0.0400      0.82150        0.978
##    12     52       2   0.8621  0.0453      0.77774        0.956
##    13     50       1   0.8448  0.0475      0.75660        0.943
##    15     49       2   0.8103  0.0515      0.71548        0.918
##    16     47       1   0.7931  0.0532      0.69541        0.905
##    17     46       1   0.7759  0.0548      0.67563        0.891
##    22     45       1   0.7586  0.0562      0.65611        0.877
##    23     44       1   0.7414  0.0575      0.63684        0.863
##    24     43       1   0.7241  0.0587      0.61778        0.849
##    25     42       1   0.7069  0.0598      0.59894        0.834
##    26     41       1   0.6897  0.0607      0.58030        0.820
##    27     40       1   0.6724  0.0616      0.56186        0.805
##    28     39       1   0.6552  0.0624      0.54359        0.790
##    30     38       4   0.5862  0.0647      0.47222        0.728
##    34     34       1   0.5690  0.0650      0.45478        0.712
##    38     33       1   0.5517  0.0653      0.43750        0.696
##    39     32       1   0.5345  0.0655      0.42036        0.680
```

```
##    40     31       1   0.5172  0.0656      0.40338        0.663
##    43     30       1   0.5000  0.0657      0.38655        0.647
##    53     29       1   0.4828  0.0656      0.36986        0.630
##    58     28       1   0.4655  0.0655      0.35332        0.613
##    63     27       1   0.4483  0.0653      0.33694        0.596
##    66     26       1   0.4310  0.0650      0.32070        0.579
##    78     25       1   0.4138  0.0647      0.30462        0.562
##    96     24       1   0.3966  0.0642      0.28868        0.545
##   114     23       1   0.3793  0.0637      0.27291        0.527
##   119     22       1   0.3621  0.0631      0.25730        0.510
##   130     21       1   0.3448  0.0624      0.24185        0.492
##   132     20       1   0.3276  0.0616      0.22657        0.474
##   141     19       1   0.3103  0.0607      0.21146        0.455
##   152     18       2   0.2759  0.0587      0.18181        0.419
##   154     16       1   0.2586  0.0575      0.16727        0.400
##   156     15       1   0.2414  0.0562      0.15295        0.381
##   177     14       1   0.2241  0.0548      0.13886        0.362
##   185     13       1   0.2069  0.0532      0.12500        0.342
##   190     12       1   0.1897  0.0515      0.11141        0.323
##   196     11       1   0.1724  0.0496      0.09811        0.303
##   201     10       1   0.1552  0.0475      0.08512        0.283
##   245      9       1   0.1379  0.0453      0.07248        0.262
##   292      8       1   0.1207  0.0428      0.06025        0.242
##   318      7       1   0.1034  0.0400      0.04849        0.221
##   333      6       1   0.0862  0.0369      0.03730        0.199
##   402      5       1   0.0690  0.0333      0.02679        0.178
##   447      4       1   0.0517  0.0291      0.01718        0.156
##   511      3       1   0.0345  0.0240      0.00883        0.135
##   536      2       1   0.0172  0.0171      0.00247        0.120
##   562      1       1   0.0000     NaN           NA           NA
```

```r
# Assuming you have a survfit object named 'survfit_obj'
summary_obj <- km_table

# Extract relevant information from the summary object
summary_data <- summary_obj$surv
lower_ci <- summary_obj$lower
upper_ci <- summary_obj$upper

# Create a data frame from the extracted information
summary_df <- data.frame(Survival = summary_data, Lower_CI = lower_ci, Upper_CI = upper_ci)

# Display the table using kable
kable(summary_df, format = "markdown", digits = 2, caption = "Survival Summary")
```

Table 1: Survival Summary

Survival	Lower_CI	Upper_CI
0.98	0.95	1.00
0.95	0.89	1.00
0.91	0.84	0.99
0.90	0.82	0.98
0.86	0.78	0.96
0.84	0.76	0.94

Survival	Lower_CI	Upper_CI
0.81	0.72	0.92
0.79	0.70	0.90
0.78	0.68	0.89
0.76	0.66	0.88
0.74	0.64	0.86
0.72	0.62	0.85
0.71	0.60	0.83
0.69	0.58	0.82
0.67	0.56	0.80
0.66	0.54	0.79
0.59	0.47	0.73
0.57	0.45	0.71
0.55	0.44	0.70
0.53	0.42	0.68
0.52	0.40	0.66
0.50	0.39	0.65
0.48	0.37	0.63
0.47	0.35	0.61
0.45	0.34	0.60
0.43	0.32	0.58
0.41	0.30	0.56
0.40	0.29	0.54
0.38	0.27	0.53
0.36	0.26	0.51
0.34	0.24	0.49
0.33	0.23	0.47
0.31	0.21	0.46
0.28	0.18	0.42
0.26	0.17	0.40
0.24	0.15	0.38
0.22	0.14	0.36
0.21	0.13	0.34
0.19	0.11	0.32
0.17	0.10	0.30
0.16	0.09	0.28
0.14	0.07	0.26
0.12	0.06	0.24
0.10	0.05	0.22
0.09	0.04	0.20
0.07	0.03	0.18
0.05	0.02	0.16
0.03	0.01	0.13
0.02	0.00	0.12
0.00	NA	NA

```r
# Access the confidence interval values from the data set
ci_lower <- km_fit$lower
ci_upper <- km_fit$upper

# Plot the survival function with 95% point wise confidence intervals
plot(km_fit, xlab = "Time", ylab = "Survival Function", main = "Kaplan-Meir Estimate of the Survival Fu
lines(km_fit, col = "blue")
```

```
lines(km_fit$time, ci_upper, type = "s", col = "red", lty = 2)
lines(km_fit$time, ci_lower, type = "s", col = "green", lty = 2)

# Add legend
legend("topright", legend = c("Survival Function", "upper","lower"),
       col = c("blue", "red","green"), lty = c(1, 2,2), bty = "n")
mtext("Figure 1: KM-Estimate", side = 1, line = 3, at = 0.5, cex = 1.2)
```

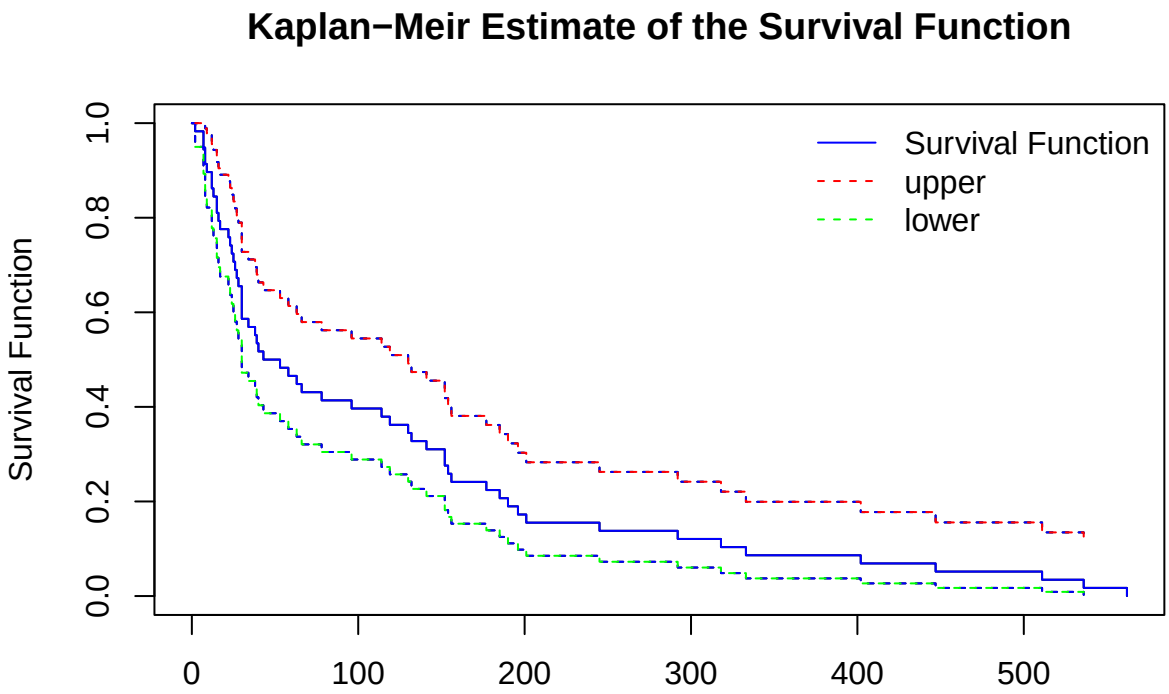

Figure 1: KM−Estimate

Estimating the Cummulative Hazard using the Nelson Aalen Estimator

#Create a survival object using the Surv() function, considering only the observed event times

```
kidney_data<-kidney
surv_obj <- Surv(time = kidney_data$time[kidney_data$status == 1], event=kidney_data$status[kidney_data$
```

#Apply the Nelson-Aalen estimator to estimate the cumulative hazard

```
na_fit <- survfit(surv_obj ~ 1, data = kidney_data, type = "fh")
```

#Extract the cumulative hazard estimates and the confidence intervals

```
na_cumhaz <- -log(na_fit$surv)
na_lower <- -log(na_fit$lower)
na_upper <- -log(na_fit$upper)
time <- na_fit$time
```

#Create a data frame with the cumulative hazard estimates and the confidence intervals

```
library(knitr)
df_nelson <- data.frame(time, na_cumhaz, na_lower, na_upper)
na_table<-kable(df_nelson)
na_table
```

time	na_cumhaz	na_lower	na_upper
2	0.0172414	0.0510339	0.0000000
7	0.0526424	0.1122179	0.0000000
8	0.0893427	0.1676786	0.0110068
9	0.1082106	0.1948367	0.0215846
12	0.1470493	0.2490377	0.0450608
13	0.1670493	0.2763114	0.0577871
15	0.2082908	0.3316012	0.0849803
16	0.2295673	0.3597383	0.0993964
17	0.2513065	0.3882733	0.1143397
22	0.2735287	0.4172539	0.1298035
23	0.2962560	0.4467257	0.1457862
24	0.3195118	0.4767337	0.1622899
25	0.3433213	0.5073226	0.1793200
26	0.3677116	0.5385379	0.1968852
27	0.3927116	0.5704264	0.2149967
28	0.4183526	0.6030366	0.2336686
30	0.5280446	0.7417605	0.3143287
34	0.5574564	0.7788102	0.3361025
38	0.5877594	0.8169428	0.3585760
39	0.6190094	0.8562360	0.3817828
40	0.6512675	0.8967747	0.4057602
43	0.6846008	0.9386522	0.4305494
53	0.7190836	0.9819711	0.4561960
58	0.7547978	1.0268450	0.4827507
63	0.7918349	1.0734004	0.5102693
66	0.8302964	1.1217785	0.5388144
78	0.8702964	1.1721376	0.5684552
96	0.9119631	1.2246567	0.5992694
114	0.9554413	1.2795386	0.6313441
119	1.0008959	1.3370148	0.6647769
130	1.0485149	1.3973512	0.6996787
132	1.0985149	1.4608551	0.7361748
141	1.1511465	1.5278846	0.7744085
152	1.2655256	1.6742800	0.8567712
154	1.3280256	1.7547408	0.9013104
156	1.3946923	1.8409646	0.9484199
177	1.4661208	1.9338369	0.9984048
185	1.5430439	2.0344590	1.0516288
190	1.6263772	2.1442243	1.1085302
196	1.7172863	2.2649297	1.1696430
201	1.8172863	2.3989458	1.2356268
245	1.9283975	2.5494878	1.3073071
292	2.0533975	2.7210620	1.3857329
318	2.1962546	2.9202526	1.4722566
333	2.3629213	3.1572010	1.5686415
402	2.5629213	3.4486631	1.6771794
447	2.8129213	3.8251613	1.8006813

time	na_cumhaz	na_lower	na_upper
511	3.1462546	4.3510195	1.9414896
536	3.6462546	5.1992592	2.0932500
562	4.6462546	7.1469109	2.1455983

Plot of the Cumulative Hazard function

```
# Plot the hazard function with 95% confidence intervals
plot(df_nelson$time, df_nelson$na_cumhaz, type = "s", col = "red",
     xlab = "Time", ylab = "Hazard", main = "Nelson-Aalen Estimate of the Cumulative Hazard")

# Add the upper and lower confidence intervals
lines(df_nelson$time, df_nelson$na_lower, type = "s", col = "purple", lty = 2)
lines(df_nelson$time, df_nelson$na_upper, type = "s", col = "green", lty = 2)
# Add a legend
legend("topleft", legend = c("Hazard Function", "Upper","Lower"),
       col = c("red","purple","green"), lty = c(1, 2,2), bty = "n")
# Add figure label
mtext("Figure 2: NA Estimator", side = 1, line = 3, at = 0.5, cex = 1.2)
```

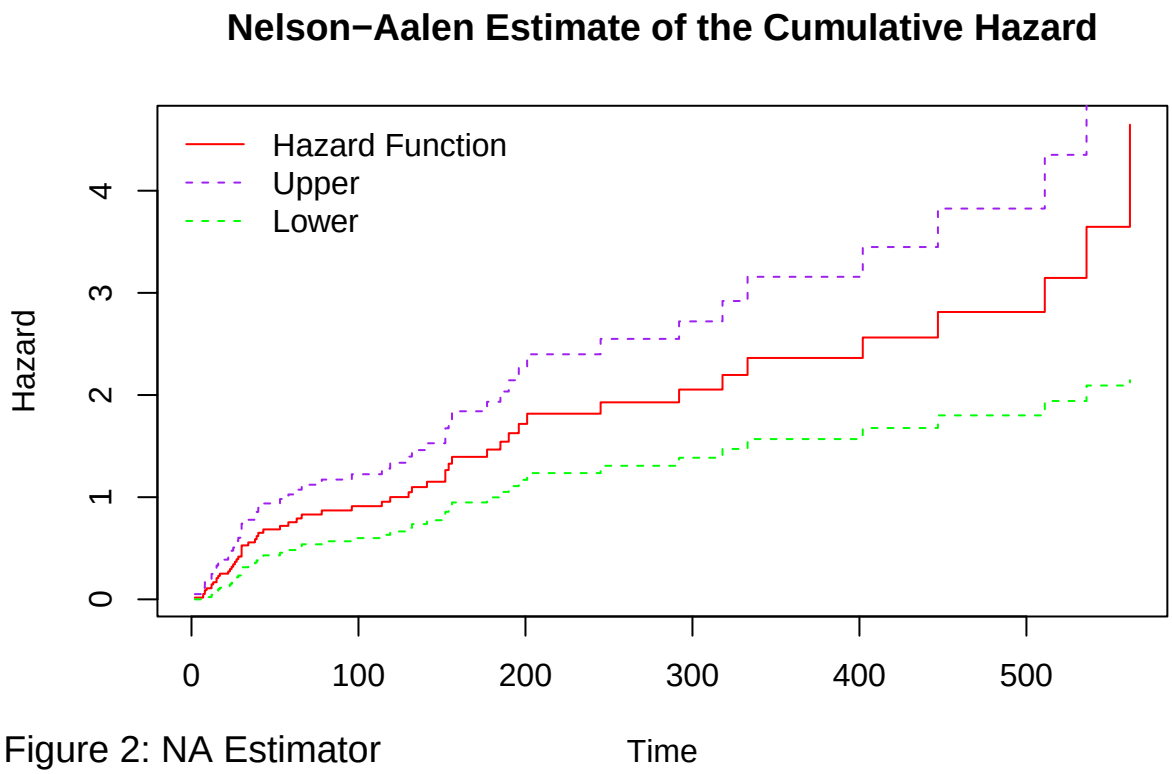

Figure 2: NA Estimator