

KAPMEIER Example #1

SUDAAN Statements and Results Illustrated

- Kaplan-Meier survival probability curve
- STRHAZ
- DESIGN
- EVENT
- TIME

Input Data Set(s): Crouchley and Pickles (1993)

Example

Consider a cross-over clinical trial with multivariate failure time data. Compute the Kaplan-Meier survival probability curve corresponding to time to onset of angina pectoris, at each level of a treatment variable.

Solution

This example demonstrates the KAPMEIER procedure in the context of correlated data from a clinical trial. The data for this example are reported by Crouchley and Pickles (1993) and represent repeated exercise times (in seconds) to angina pectoris in patients with coronary heart disease. The event of interest is onset of angina pectoris, and we use the terms “onset times”, “event times”, and “survival times” to mean the same thing.

Briefly, a total of 21 subjects were each tested four times on one day and a further four times two days later. On each day, exercise time measurements were taken at 4 times: just before dosing and at 1 hour, 3 hours, and 5 hours following drug administration. On Day 2, the drug was an active treatment (an oral dose of isosorbide dinitrate), and on Day 1 a placebo. The dataset consists of 168 records—21 patients and 8 records per patient (4 records/day/patient).

The Kaplan-Meier survival probability curve corresponding to the onset time to angina pectoris was estimated for each level of the treatment variable (Day 1 Placebo vs. Day 2 Treatment) at each of the 4 times—pre-dosing, 1 hr, 3 hrs, and 5 hrs post-dosing. This was accomplished by specifying both the treatment and hours variables on the STRHAZ statement.

The following SUDAAN program contains code to fit the Kaplan-Meier curve to the observed event times. The default sample design option DESIGN=WR (shorthand notation for “with replacement sampling”) invokes the robust variance estimator that is appropriate for the study. The NEST statement in SUDAAN indicates that the patient (PATIENT) represents the cluster or primary sampling unit, with the keyword `_ONE_` indicating there is a single sample design stratum. Additional sources of intracluster correlation, such as time within day, need not be specified.

The program was run in SAS-Callable SUDAAN, and the associated program code and results are provided below.

Exhibit 5. SAS-Callable SUDAAN Code for KAPMEIER

```
libname in "c:\11winbetatest\examples";

proc format;
  value hrs 1="1 hr"
           2="3 hrs"
           3="5 hrs"
           4="Pre-Dosing";
  value trt 1="Placebo Day"
           2="Treatment Day";

PROC KAPMEIER DATA=in.exercise DESIGN=WR FILETYPE=SAS;
  NEST _ONE_ PATIENT;
  WEIGHT _ONE_;

  EVENT COMPLETE;
  TIME EXTIME ;
  CLASS HRS TRTDAY;
  STRHAZ HRS TRTDAY;

  RFORMAT trtday trt.;
  RFORMAT hrs hrs.;
  SETENV COLWIDTH=8;
  OUTPUT / KAPMEIER=all FILENAME="C:\11winbetatest\Examples\KM_HRS_OUT"
          FILETYPE=SUDAAN REPLACE;
  RTITLE "Kaplan-Meier Survival Probabilities for Exercise Times to Angina Pectoris:
  At Event Times"
  "Stratified by HOURS (1=1hr, 2=3hrs, 3=5hrs, 4=Pre-Dosing) and
  DAY (1=Placebo, 2=Treatment)";
  RFOOTNOTE "Source Crouchley and Pickles (1993, Biometrics 49, 1067-1076)";

PROC RECORDS DATA="C:\11winbetatest\Examples\KM_HRS_OUT" FILETYPE=SUDAAN contents;
  SETENV COLWIDTH=6 decwidth=4;
  RLABEL hrs="Hours";
  RLABEL sekm="SE";
  RLABEL lowkm="Lower 95% Limit";
  RLABEL upkm="Upper 95% Limit";
  PRINT hrs trtday extime km sekm lowkm upkm /
        extimefmt=f8.0 kmfmt=f8.4 trtdayfmt=f6.0 hrsfmt=f6.0 sekmfmt=f7.4;
  RTITLE "Printing the KM Curves";
```

Exhibit 5 contains the SAS-Callable SUDAAN code for this example. The KAPMEIER procedure generates the Kaplan-Meier estimates at the unique event times of the variable EXTIME (exercise times to onset of event), within levels of treatment day (TRTDAY) within hours (HRS). The censoring variable COMPLETE=1 if the exercise time is a complete time (event occurred), 0 if a censored time (event did not occur).

The OUTPUT statement directs the Kaplan-Meier estimates to an output data set for additional processing. A SUDAAN dataset was created, but other types of datasets could also have been created (e.g., SAS dataset).

The RECORDS procedure prints the output dataset containing the Kaplan-Meier estimates. Some labeling and formatting was included, but this is not required. Even the PRINT statement could have been excluded. RLABEL statements were used to shorten the default labels of some of the PRINT keywords. Alternatively, the PRINT statement could have been used to do the same relabeling:

```
PRINT hrs="Hours" trtday extime km sekm="SE" lowkm="Lower 95% Limit"
      upkm="Upper 95% Limit" / extimefmt=f8.0 kmfmt=f8.4 trtdayfmt=f6.0
      hrsfmt=f6.0 sekmfmt=f7.4;
```

Exhibit 6. First Page of KAPMEIER Results (*.lst file)

```

                                S U D A A N
Software for the Statistical Analysis of Correlated Data
Copyright      Research Triangle Institute      December 2011
                                Release 11.0.0

DESIGN SUMMARY: Variances will be computed using the Taylor Linearization Method, Assuming a
With Replacement (WR) Design
  Sample Weight:  _ONE_
  Stratification Variables(s):  _ONE_
  Primary Sampling Unit:  PATIENT

Number of observations read      :   168      Weighted count:      168
Observations used in the analysis :   168      Weighted count:      168
Denominator degrees of freedom   :    20

```

Exhibit 6 through *Exhibit 10* contain the limited printed results from KAPMEIER. *Exhibit 6* displays the number of observations read (168), used in analysis (168), and denominator degrees of freedom (21-1=20),

Exhibit 7. Frequency of Censored Data within Strata

```

Summary of Event Values
by: COMPLETE, Hours Since Drug Admin, Day.

for: COMPLETE = Censored.
-----
Hours Since Drug Admin
  Day              Frequency      Weighted Sum
-----
1 hr
  Placebo Day      0.000          0.000
  Treatment Day    7.000          7.000
3 hrs
  Placebo Day      0.000          0.000
  Treatment Day    4.000          4.000
5 hrs
  Placebo Day      0.000          0.000
  Treatment Day    2.000          2.000
Pre-Dosing
  Placebo Day      0.000          0.000
  Treatment Day    0.000          0.000
-----

```

Exhibit 7 contains the number of censored event times within levels of stratification variables—hours since drug administration and treatment day. Censoring occurs when no event is experienced during the observation period. We see that there are no censored times during pre-dosing on either day—*i.e.*, all subjects experience the event during the observation period without treatment. At 1 hr, 3 hrs, and 5 hrs on placebo (Day 1), the same is true. However, at 1 hr, 3 hrs, and 5 hrs post treatment (Day 2), some patients do not experience the event during the observation period—greatest at 1 hr (7 patients with no event), but also 4 patients at 3 hrs and 2 patients at 5 hrs.

Exhibit 8. Frequency of Non-Censored Data within Strata

Summary of Event Values
 by: COMPLETE, Hours Since Drug Admin, Day.

for: COMPLETE = **Non-Censored**.

Hours Since Drug Admin Day	Frequency	Weighted Sum
1 hr		
Placebo Day	21.000	21.000
Treatment Day	14.000	14.000
3 hrs		
Placebo Day	21.000	21.000
Treatment Day	17.000	17.000
5 hrs		
Placebo Day	21.000	21.000
Treatment Day	19.000	19.000
Pre-Dosing		
Placebo Day	21.000	21.000
Treatment Day	21.000	21.000

Exhibit 8 provides the frequencies for non-censored (*i.e.*, complete) event times. This table is just the complement to the one previous. Pre-dosing on each day and all time points on the placebo day contain only event times, while there are fewer patients who experience the event at 1, 3, and 5 hours post-dosing on the treatment day.

Exhibit 9. Frequencies for CLASS Variable HRS

Frequencies and Values for CLASS Variables

Hours Since Drug Admin	Frequency	Value
Ordered Position:		
1	42	1 hr
Ordered Position:		
2	42	3 hrs
Ordered Position:		
3	42	5 hrs
Ordered Position:		
4	42	Pre-Dosing

Exhibit 10. Frequencies for CLASS Variable TRTDAY

Frequencies and Values for CLASS Variables		
Day	Frequency	Value
Ordered		
Position:		
1	84	Placebo Day
Ordered		
Position:		
2	84	Treatment Day

Exhibit 9 and *Exhibit 10* contain the frequency distributions for CLASS variables HRS and TRTDAY. This ends the printed results from the KAPMEIER procedure.

The OUTPUT statement in KAPMEIER requested that the KM_HRS_OUT.SDN data set be created and contain all of the pertinent statistics. After the data set was created, PROC RECORDS was used to print this data set. The results from PROC RECORDS are presented next.

Exhibit 11 contains the first page of results from PROC RECORDS. This is a display of contents and value labels on the output dataset created by KAPMEIER: KM_HRS_OUT.SDN.

Exhibit 11. First Page of RECORDS Results (*.lst file)

```

          S U D A A N
    Software for the Statistical Analysis of Correlated Data
  Copyright      Research Triangle Institute      December 2011
                Release 11.0.0

SUDAAN Release 11.00
Record file C:\11winbetatest\Examples\KM_HRS_OUT.SDN created 12-20-2011 11:31:43
Variables
Name          Type          Output      Description
                Format
-----
PROCNUM       Numeric      F2.0        Procedure Number
EXTIME        Numeric      F8.2        Exercise Time to Angina Pectoris
KM            Numeric      F10.2       Kaplan-Meier Estimate
SEKM         Numeric      F10.2       SE Kaplan-Meier Est.
LOWKM        Numeric      F10.2       Lower 95% Confidence Limit
UPKM         Numeric      F10.2       Upper 95% Confidence Limit
HRS          Numeric      F8.2        Hours Since Drug Admin
TRTDAY       Numeric      F8.2        Day

Values and Labels for Variable PROCNUM:
Value      Label
-----
14      KAPMEIER

Values and Labels for Variable HRS:
Value      Label
-----
1      1 hr
2      3 hrs
3      5 hrs
4      Pre-Dosing

Values and Labels for Variable TRTDAY:
Value      Label
-----
1      Placebo Day
2      Treatment Day

```

The printed dataset containing the Kaplan-Meier estimates and related statistics is contained in **Exhibit 12** (it spans multiple pages).

Exhibit 12. Kaplan-Meier Survival Probability Estimates (By HRS and TRTDAY)

Variance Estimation Method: Taylor Series (WR)
 Time Variable: EXTIME: Exercise Time to Angina Pectoris
 Censoring Variable: COMPLETE

Kaplan-Meier Survival Probabilities for Exercise Times to Angina Pectoris: At Event Times
 Stratified by HOURS (1=1hr, 2=3hrs, 3=5hrs, 4=Pre-Dosing) and DAY (1=Placebo, 2=Treatment)
 Printing the KM Curves

by: Observation Number.

Observation Number	Hours	Day	Exercise Time to Angina Pectoris	Kaplan- Meier Estimate	SE	Lower 95% Limit	Upper 95% Limit
1	1	1	114	0.9524	0.0476	0.6612	0.9943
2	1	1	125	0.9048	0.0656	0.6351	0.9782
3	1	1	157	0.8571	0.0782	0.5885	0.9562
4	1	1	168	0.8095	0.0878	0.5398	0.9301
5	1	1	172	0.7619	0.0952	0.4919	0.9010
6	1	1	194	0.7143	0.1010	0.4455	0.8693
7	1	1	199	0.6190	0.1086	0.3575	0.7996
8	1	1	210	0.5714	0.1107	0.3161	0.7619
9	1	1	224	0.5238	0.1117	0.2763	0.7225
10	1	1	232	0.4762	0.1117	0.2382	0.6813
11	1	1	244	0.4286	0.1107	0.2019	0.6384
12	1	1	249	0.3810	0.1086	0.1675	0.5938
13	1	1	264	0.3333	0.1054	0.1350	0.5473
14	1	1	287	0.2857	0.1010	0.1047	0.4989
15	1	1	290	0.2381	0.0952	0.0768	0.4483
16	1	1	329	0.1905	0.0878	0.0518	0.3951
17	1	1	357	0.1429	0.0782	0.0302	0.3390
18	1	1	390	0.0952	0.0656	0.0131	0.2792
19	1	1	431	0.0476	0.0476	0.0024	0.2156
20	1	1	512	0.0000	0.0000	.	.
21	1	2	110	0.9524	0.0476	0.6612	0.9943
22	1	2	121	0.9048	0.0656	0.6351	0.9782
23	1	2	145	0.8571	0.0782	0.5885	0.9562
24	1	2	230	0.8095	0.0878	0.5398	0.9301
25	1	2	232	0.7619	0.0952	0.4919	0.9010
26	1	2	248	0.7143	0.1010	0.4455	0.8693
27	1	2	250	0.6667	0.1054	0.4007	0.8355
28	1	2	264	0.6190	0.1086	0.3575	0.7996
29	1	2	306	0.5714	0.1107	0.3161	0.7619
30	1	2	327	0.5238	0.1117	0.2763	0.7225
31	1	2	403	0.4762	0.1117	0.2382	0.6813
32	1	2	432	0.4286	0.1107	0.2019	0.6384
33	1	2	580	0.3214	0.1262	0.0968	0.5761
34	1	2	651	0.2143	0.1229	0.0351	0.4925
35	2	1	99	0.9524	0.0476	0.6612	0.9943
36	2	1	118	0.9048	0.0656	0.6351	0.9782
37	2	1	135	0.8571	0.0782	0.5885	0.9562
38	2	1	136	0.8095	0.0878	0.5398	0.9301
39	2	1	147	0.7619	0.0952	0.4919	0.9010
40	2	1	168	0.7143	0.1010	0.4455	0.8693
41	2	1	188	0.6667	0.1054	0.4007	0.8355
42	2	1	197	0.5714	0.1107	0.3161	0.7619
43	2	1	205	0.5238	0.1117	0.2763	0.7225
44	2	1	211	0.4762	0.1117	0.2382	0.6813

Source Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

Exhibit 12. Kaplan-Meier Survival Probability Estimates (By HRS and TRTDAY) (cont'd)

Variance Estimation Method: Taylor Series (WR)
 Time Variable: EXTIME: Exercise Time to Angina Pectoris
 Censoring Variable: COMPLETE

Kaplan-Meier Survival Probabilities for Exercise Times to Angina Pectoris: At Event Times
 Stratified by HOURS (1=1hr, 2=3hrs, 3=5hrs, 4=Pre-Dosing) and DAY (1=Placebo, 2=Treatment)
 Printing the KM Curves

by: Observation Number

Observation Number	Hours	Day	Exercise Time to Angina Pectoris	Kaplan-Meier Estimate	SE	Lower 95% Limit	Upper 95% Limit
45	2	1	223	0.4286	0.1107	0.2019	0.6384
46	2	1	243	0.3810	0.1086	0.1675	0.5938
47	2	1	251	0.3333	0.1054	0.1350	0.5473
48	2	1	267	0.2857	0.1010	0.1047	0.4989
49	2	1	276	0.2381	0.0952	0.0768	0.4483
50	2	1	281	0.1905	0.0878	0.0518	0.3951
51	2	1	388	0.1429	0.0782	0.0302	0.3390
52	2	1	441	0.0952	0.0656	0.0131	0.2792
53	2	1	448	0.0476	0.0476	0.0024	0.2156
54	2	1	560	0.0000	0.0000	.	.
55	2	2	110	0.9524	0.0476	0.6612	0.9943
56	2	2	123	0.9048	0.0656	0.6351	0.9782
57	2	2	150	0.8571	0.0782	0.5885	0.9562
58	2	2	172	0.8095	0.0878	0.5398	0.9301
59	2	2	206	0.7619	0.0952	0.4919	0.9010
60	2	2	210	0.7143	0.1010	0.4455	0.8693
61	2	2	258	0.6667	0.1054	0.4007	0.8355
62	2	2	264	0.6190	0.1086	0.3575	0.7996
63	2	2	280	0.5714	0.1107	0.3161	0.7619
64	2	2	290	0.5238	0.1117	0.2763	0.7225
65	2	2	291	0.4762	0.1117	0.2382	0.6813
66	2	2	298	0.4286	0.1107	0.2019	0.6384
67	2	2	370	0.3810	0.1086	0.1675	0.5938
68	2	2	492	0.3265	0.1064	0.1281	0.5435
69	2	2	566	0.2177	0.1154	0.0429	0.4781
70	2	2	613	0.1088	0.0977	0.0057	0.3855
71	2	2	624	0.0000	0.0000	.	.
72	3	1	105	0.9524	0.0476	0.6612	0.9943
73	3	1	111	0.9048	0.0656	0.6351	0.9782
74	3	1	115	0.8571	0.0782	0.5885	0.9562
75	3	1	143	0.8095	0.0878	0.5398	0.9301
76	3	1	159	0.7619	0.0952	0.4919	0.9010
77	3	1	176	0.7143	0.1010	0.4455	0.8693
78	3	1	189	0.6667	0.1054	0.4007	0.8355
79	3	1	206	0.6190	0.1086	0.3575	0.7996
80	3	1	207	0.5714	0.1107	0.3161	0.7619
81	3	1	210	0.5238	0.1117	0.2763	0.7225
82	3	1	222	0.4762	0.1117	0.2382	0.6813
83	3	1	223	0.4286	0.1107	0.2019	0.6384
84	3	1	227	0.3810	0.1086	0.1675	0.5938
85	3	1	237	0.3333	0.1054	0.1350	0.5473
86	3	1	241	0.2857	0.1010	0.1047	0.4989
87	3	1	250	0.2381	0.0952	0.0768	0.4483
88	3	1	251	0.1905	0.0878	0.0518	0.3951

Source Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

Exhibit 12. Kaplan-Meier Survival Probability Estimates (By HRS and TRTDAY) (cont'd)

Variance Estimation Method: Taylor Series (WR)
Time Variable: EXTIME: Exercise Time to Angina Pectoris
Censoring Variable: COMPLETE

Kaplan-Meier Survival Probabilities for Exercise Times to Angina Pectoris: At Event Times
Stratified by HOURS (1=1hr, 2=3hrs, 3=5hrs, 4=Pre-Dosing) and DAY (1=Placebo, 2=Treatment)
Printing the KM Curves

by: Observation Number.

Observation Number	Hours	Day	Exercise Time to Angina Pectoris	Kaplan- Meier Estimate	SE	Lower 95% Limit	Upper 95% Limit
89	3	1	328	0.1429	0.0782	0.0302	0.3390
90	3	1	388	0.0952	0.0656	0.0131	0.2792
91	3	1	468	0.0476	0.0476	0.0024	0.2156
92	3	1	478	0.0000	0.0000	.	.
93	3	2	102	0.9524	0.0476	0.6612	0.9943
94	3	2	105	0.9048	0.0656	0.6351	0.9782
95	3	2	123	0.8571	0.0782	0.5885	0.9562
96	3	2	172	0.8095	0.0878	0.5398	0.9301
97	3	2	180	0.7619	0.0952	0.4919	0.9010
98	3	2	207	0.7143	0.1010	0.4455	0.8693
99	3	2	208	0.6667	0.1054	0.4007	0.8355
100	3	2	212	0.6190	0.1086	0.3575	0.7996
101	3	2	216	0.5714	0.1107	0.3161	0.7619
102	3	2	224	0.5238	0.1117	0.2763	0.7225
103	3	2	226	0.4762	0.1117	0.2382	0.6813
104	3	2	268	0.4286	0.1107	0.2019	0.6384
105	3	2	303	0.3810	0.1086	0.1675	0.5938
106	3	2	316	0.3333	0.1054	0.1350	0.5473
107	3	2	321	0.2857	0.1010	0.1047	0.4989
108	3	2	391	0.2381	0.0952	0.0768	0.4483
109	3	2	439	0.1905	0.0878	0.0518	0.3951
110	3	2	514	0.1270	0.0790	0.0208	0.3329
111	3	2	554	0.0000	0.0000	.	.
112	4	1	87	0.9524	0.0476	0.6612	0.9943
113	4	1	108	0.9048	0.0656	0.6351	0.9782
114	4	1	131	0.8571	0.0782	0.5885	0.9562
115	4	1	150	0.7619	0.0952	0.4919	0.9010
116	4	1	190	0.7143	0.1010	0.4455	0.8693
117	4	1	194	0.6667	0.1054	0.4007	0.8355
118	4	1	199	0.6190	0.1086	0.3575	0.7996
119	4	1	205	0.5714	0.1107	0.3161	0.7619
120	4	1	215	0.5238	0.1117	0.2763	0.7225
121	4	1	218	0.4762	0.1117	0.2382	0.6813
122	4	1	221	0.4286	0.1107	0.2019	0.6384
123	4	1	227	0.3810	0.1086	0.1675	0.5938
124	4	1	228	0.3333	0.1054	0.1350	0.5473
125	4	1	234	0.2381	0.0952	0.0768	0.4483
126	4	1	277	0.1905	0.0878	0.0518	0.3951
127	4	1	301	0.1429	0.0782	0.0302	0.3390
128	4	1	342	0.0952	0.0656	0.0131	0.2792
129	4	1	406	0.0476	0.0476	0.0024	0.2156
130	4	1	424	0.0000	0.0000	.	.
131	4	2	89	0.9524	0.0476	0.6612	0.9943
132	4	2	129	0.9048	0.0656	0.6351	0.9782

Source Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

Exhibit 12. Kaplan-Meier Survival Probability Estimates (By HRS and TRTDAY) (cont'd)

Variance Estimation Method: Taylor Series (WR)
 Time Variable: EXTIME: Exercise Time to Angina Pectoris
 Censoring Variable: COMPLETE

Kaplan-Meier Survival Probabilities for Exercise Times to Angina Pectoris: At Event Times
 Stratified by HOURS (1=1hr, 2=3hrs, 3=5hrs, 4=Pre-Dosing) and DAY (1=Placebo, 2=Treatment)
 Printing the KM Curves

by: Observation Number.

Observation Number	Hours	Day	Exercise Time to Angina Pectoris	Kaplan-Meier Estimate	SE	Lower 95% Limit	Upper 95% Limit
133	4	2	136	0.8571	0.0782	0.5885	0.9562
134	4	2	147	0.8095	0.0878	0.5398	0.9301
135	4	2	152	0.7619	0.0952	0.4919	0.9010
136	4	2	154	0.7143	0.1010	0.4455	0.8693
137	4	2	208	0.6667	0.1054	0.4007	0.8355
138	4	2	213	0.6190	0.1086	0.3575	0.7996
139	4	2	215	0.5714	0.1107	0.3161	0.7619
140	4	2	224	0.5238	0.1117	0.2763	0.7225
141	4	2	229	0.4762	0.1117	0.2382	0.6813
142	4	2	231	0.4286	0.1107	0.2019	0.6384
143	4	2	235	0.3810	0.1086	0.1675	0.5938
144	4	2	250	0.2857	0.1010	0.1047	0.4989
145	4	2	265	0.2381	0.0952	0.0768	0.4483
146	4	2	406	0.1905	0.0878	0.0518	0.3951
147	4	2	417	0.1429	0.0782	0.0302	0.3390
148	4	2	425	0.0952	0.0656	0.0131	0.2792
149	4	2	441	0.0476	0.0476	0.0024	0.2156
150	4	2	490	0.0000	0.0000	.	.

Source Crouchley and Pickles (1993, Biometrics 49, 1067-1076)

Exhibit 12 contains the estimated Kaplan-Meier survival probability curve (with SEs and 95% CIs) at each combination of hours since drug administration and treatment day. Note that within any combination of hours and day, only unique event times are represented. All of the observations are used in the calculation of the Kaplan-Meier estimates, but the corresponding survival probability curves are step functions that only depend on knowing the survival probabilities at unique event times.

These estimates could be used to compare median survival time (time at which survival probability is 50%) for each day (treatment vs. placebo) and hours since drug administration or to plot the Kaplan-Meier survival probability curves at each combination of day and hours.