

## Estimation for exponential and Weibull distributions

```
-----
log: C:\Documents and Settings\Qing\sur641Lab2.log
log type: text
opened on: 1 Mar 2005, 12:49:21

#####
##Some preparation about the data
#####
. pwd
C:\Documents and Settings\Qing

. cd ..
C:\Documents and Settings

. mkdir sur641Lab2

. cd sur641Lab2
C:\Documents and Settings\sur641Lab2

. insheet using survivalOneSample.txt
you must start with an empty dataset
r(18);

. clear

. insheet using survivalOneSample.txt
(3 vars, 40 obs)

. edit
- preserve

. clear

#####
##Now, we have our data, do some data exploration
#####
. use "C:\Documents and
Settings\sur641Lab2\survivalOneSample.dta", clear

. summarize

Variable | Obs Mean Std. Dev. Min Max
-----|-----
t | 40 2.675 2.197289 .1 7.9
cc | 40 .4 .4961389 0 1

. summarize, detail

t
-----
Percentiles Smallest
1% .1 .1
5% .15 .1
10% .35 .2
Obs 40
```

```

25%      .9      .3      Sum of Wgt.      40
50%     2.1      Largest      Mean      2.675
75%     3.95     7.2      Std. Dev.  2.197289
90%     6.1     7.3      Variance   4.828077
95%     7.4     7.5      Skewness   .8840067
99%     7.9     7.9      Kurtosis   2.90674

```

```

-----
cc
-----
Percentiles      Smallest
1%                0                0
5%                0                0
10%               0                0      Obs      40
25%               0                0      Sum of Wgt.  40

50%               0                Largest      Mean      .4
75%               1                1          Std. Dev.  .4961389
90%               1                1          Variance   .2461538
95%               1                1          Skewness   .4082483
99%               1                1          Kurtosis   1.166667

```

```
. tabstat t, by(cc)
```

```
Summary for variables: t
by categories of: cc
```

```

cc |      mean
-----+-----
0  |  2.354167
1  |  3.15625
-----+-----
Total |      2.675
-----

```

```
. tabstat, stats(count mean range)
varlist required
r(100);
```

```
. tabstat t c, stats(count mean)
```

```

stats |      t      cc
-----+-----
N     |      40     40
mean  |      2.675  .4
-----

```

```
#####
##Tell STATA our data is survival data
#####
```

```
. stset cc, failure(t)
```

```

failure event:  t != 0 & t < .
obs. time interval:  (0, cc]
exit on or before:  failure

```

```
-----
40 total obs.
24 obs. end on or before enter()
-----
```

```

16 obs. remaining, representing
16 failures in single record/single failure data
16 total analysis time at risk, at risk from t = 0
    earliest observed entry t = 0
    last observed exit t = 1

```

**. stdes**

```

failure _d: cc
analysis time _t: t

```

Category	total	per subject			
		mean	min	median	max
no. of subjects	40				
no. of records	40	1	1	1	1
(first) entry time		0	0	0	0
(final) exit time		2.675	.1	2.1	7.9
subjects with gap	0				
time on gap if gap	0				
time at risk	107	2.675	.1	2.1	7.9
failures	16	.4	0	0	1

**. stsum**

```

failure _d: cc
analysis time _t: t

```

	time at risk	incidence rate	no. of subjects	Survival time		
				25%	50%	75%
total	106.9999999	.1495327	40	2.7	4.5	7.9

**. sts list**

```

failure _d: cc
analysis time _t: t

```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
.1	40	0	2	1.0000	.	.	.
.2	38	0	1	1.0000	.	.	.
.3	37	0	1	1.0000	.	.	.
.4	36	0	1	1.0000	.	.	.
2.1	21	1	1	0.8284	0.0708	0.6324	0.9256
Omitting part of the table							
4.5	8	1	0	0.4382	0.1111	0.2217	0.6365
5	7	2	1	0.3130	0.1091	0.1236	0.5246
7.2	4	0	1	0.3130	0.1091	0.1236	0.5246
7.3	3	0	1	0.3130	0.1091	0.1236	0.5246
7.5	2	0	1	0.3130	0.1091	0.1236	0.5246
7.9	1	1	0	0.0000	.	.	.

```

#####
##Can we see some graph? Please
#####

```

**. sts graph**

```

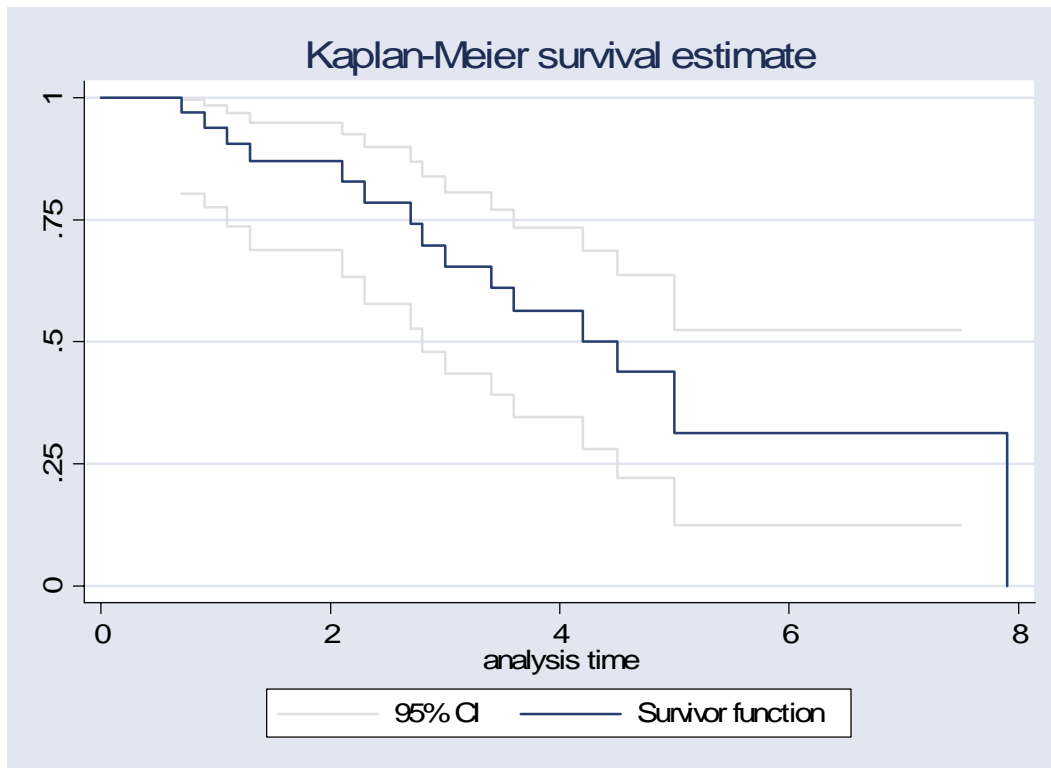
failure _d: cc
analysis time _t: t

```

```
. sts gen lb=lb(s) ub=ub(s)
```

```
. sts graph, gwood
```

```
      failure _d: cc  
analysis time _t: t
```



```
#####
##Regressions: Weibull and exponential
#####
. streg , distribution(weibull)
```

```
      failure _d: cc
      analysis time _t: t
```

Fitting constant-only model:

Weibull regression -- log relative-hazard form

```
No. of subjects =      40                Number of obs   =      40
No. of failures =      16
Time at risk   = 106.9999999
Log likelihood = -28.051073                LR chi2(0)       =      0.00
                                           Prob > chi2      =      .
```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
/ln_p	.5207668	.1886855	2.76	0.006	.1509499	.8905836
p	1.683318	.3176177			1.162938	2.436551
1/p	.5940649	.1120914			.4104162	.8598907

```
. streg , distribution(weibull) robust
```

_t	Haz. Ratio	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
/ln_p	.5207668	.1620307	3.21	0.001	.2031924	.8383411
p	1.683318	.2727492			1.225308	2.312528
1/p	.5940649	.0962568			.4324273	.8161212

```
. streg , distribution(exponential) nohr
```

```
failure _d: cc  
analysis time _t: t
```

```
Iteration 0: log likelihood = -31.082263  
Iteration 1: log likelihood = -31.082263
```

```
Exponential regression -- log relative-hazard form
```

```
No. of subjects =          40          Number of obs   =          40  
No. of failures =          16  
Time at risk   = 106.9999999  
Log likelihood = -31.082263          LR chi2(0)       =          -0.00  
                                          Prob > chi2     =          .
```

```
-----+-----  
_t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]  
-----+-----  
_cons | -1.90024     .25     -7.60   0.000    -2.390231   -1.410249  
-----+-----
```

**Estimation for exponential and Weibull two-sample models**

```
#####
##Exploratory analysis of two sample data
#####
. use survivalTwoSample
```

```
. stset t, failure(cc)
```

```
failure event: cc != 0 & cc < .
obs. time interval: (0, t]
exit on or before: failure
```

```
-----
40 total obs.
0 exclusions
-----
40 obs. remaining, representing
16 failures in single record/single failure data
107 total analysis time at risk, at risk from t = 0
earliest observed entry t = 0
last observed exit t = 7.9
```

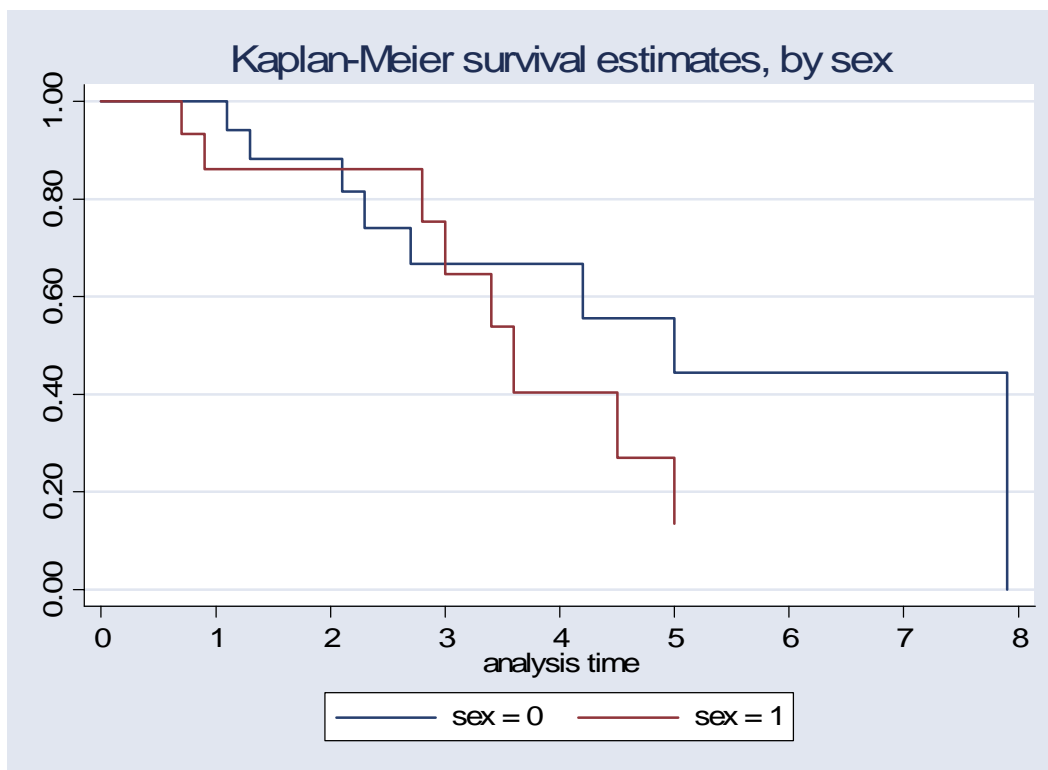
```
. sts list, by(sex)
```

```
failure _d: cc
analysis time _t: t
```

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
-----							
sex=0							
.3	21	0	1	1.0000	.	.	.
.4	20	0	1	1.0000	.	.	.
.5	19	0	1	1.0000	.	.	.
.9	18	0	1	1.0000	.	.	.
1.1	17	1	0	0.9412	0.0571	0.6502	0.9915
1.3	16	1	0	0.8824	0.0781	0.6060	0.9692
7.2	4	0	1	0.4443	0.1525	0.1558	0.7018
7.3	3	0	1	0.4443	0.1525	0.1558	0.7018
7.5	2	0	1	0.4443	0.1525	0.1558	0.7018
7.9	1	1	0	0.0000	.	.	.
sex=1							
.1	19	0	2	1.0000	.	.	.
.2	17	0	1	1.0000	.	.	.
.6	16	0	1	1.0000	.	.	.
.7	15	1	0	0.9333	0.0644	0.6126	0.9903
3.6	4	1	0	0.4038	0.1663	0.1099	0.6891
4.5	3	1	0	0.2692	0.1561	0.0442	0.5758
5	2	1	1	0.1346	0.1231	0.0074	0.4402
-----							

```
. sts graph, by(sex) ylabel(0(.2)1) xlabel(0(.1)8)
```

```
failure _d: cc
analysis time _t: t
```



```
#####
. - label define gender 0 "female"- label define gender 1 "male",
add

. label values sex gender

. sts graph, by(sex) ylabel(0(.2)1) xlabel(0(1)8)

        failure _d: cc
analysis time _t: t

#####
##Two sample test
#####

. sts test sex, logrank

        failure _d: cc
analysis time _t: t

Log-rank test for equality of survivor functions

sex |      Events      Events
    | observed      expected
-----+-----
female |      8      10.21
male   |      8       5.79
-----+-----
Total  |     16      16.00

        chi2(1) =      1.42
        Pr>chi2 =      0.2331

. sts test sex, cox
```

```
failure _d: cc
analysis time _t: t
```

Cox regression-based test for equality of survival curves

sex	Events observed	Events expected	Relative hazard
female	8	10.21	0.8028
male	8	5.79	1.4731
Total	16	16.00	1.0000

LR chi2(1) = 1.35  
Pr>chi2 = 0.2449

```
#####
##Exponentail and Weibull regression
#####
. streg sex, distribution(exponential) nohr
```

```
failure _d: cc
analysis time _t: t
```

```
Iteration 0: log likelihood = -31.082263
Iteration 1: log likelihood = -30.519924
Iteration 2: log likelihood = -30.506355
Iteration 3: log likelihood = -30.506347
Iteration 4: log likelihood = -30.506347
```

Exponential regression -- log relative-hazard form

```
No. of subjects = 40 Number of obs = 40
No. of failures = 16
Time at risk = 106.9999999
Log likelihood = -30.506347 LR chi2(1) = 1.15
Prob > chi2 = 0.2832
```

_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
sex	.5398422	.5	1.08	0.280	-.4401398 1.519824
_cons	-2.134166	.3535534	-6.04	0.000	-2.827118 -1.441215

```
. streg sex, distribution(weibull) nohr
```

```
failure _d: cc
analysis time _t: t
```

Fitting constant-only model:

```
Iteration 0: log likelihood = -31.082263
Iteration 1: log likelihood = -28.134897
Iteration 2: log likelihood = -28.051129
Iteration 3: log likelihood = -28.051073
Iteration 4: log likelihood = -28.051073
```

Fitting full model:

```
Iteration 0: log likelihood = -28.051073
Iteration 1: log likelihood = -26.7777
Iteration 2: log likelihood = -26.67501
Iteration 3: log likelihood = -26.674714
Iteration 4: log likelihood = -26.674714
```

Weibull regression -- log relative-hazard form

```
No. of subjects = 40 Number of obs = 40
No. of failures = 16
Time at risk = 106.9999999
```

Log likelihood = -26.674714

LR chi2(1) = 2.75  
Prob > chi2 = 0.0971

_t	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
sex	.8737798	.5216051	1.68	0.094	-.1485475	1.896107
_cons	-3.463664	.7045342	-4.92	0.000	-4.844526	-2.082802
/ln_p	.6053631	.1942883	3.12	0.002	.224565	.9861613
p	1.831917	.3559202			1.251778	2.680923
1/p	.5458762	.1060574			.3730058	.7988637