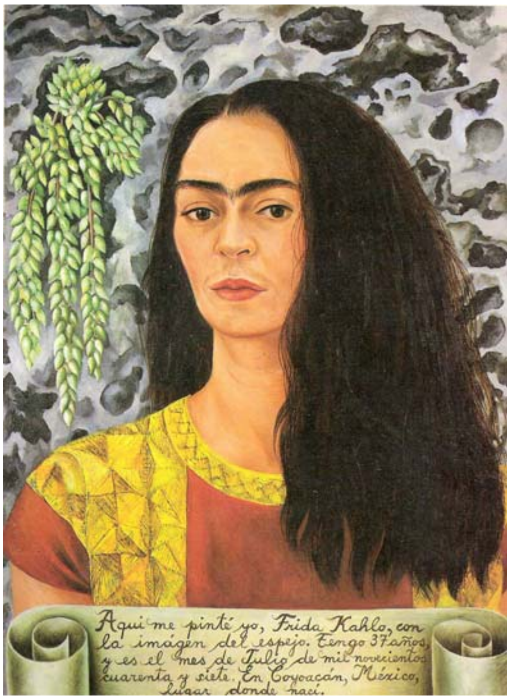
Outline of the talk

- Confidence Intervals and confidence bands of the survival function
- Validation of the estimates and examples
- Comparing Methods and Transformations
- Coverage probabilities
- Conclusions



Confidence Intervals and Confidence Bands

- The Kaplan-Meyer method is a standard estimator of the survival function, i.e. of the survival probabilities along the analysis time.
- Confidence intervals are usually derived by transformation of the survival function on the log-minus-log scale followed by the estimation of appropriate variance.
- So, let

$$\sigma = \sqrt{\sum_{t_i \le t} \frac{d_i}{n_i(n_i - d_i)}}$$

(the sum in the Greenwood's formula)

confidence intervals for the survival function are then computed as follows:

$$S(t)^{\exp\left(\pm \frac{z_{1-\alpha/2}\sigma}{\ln\left[S(t)\right]}\right)}$$

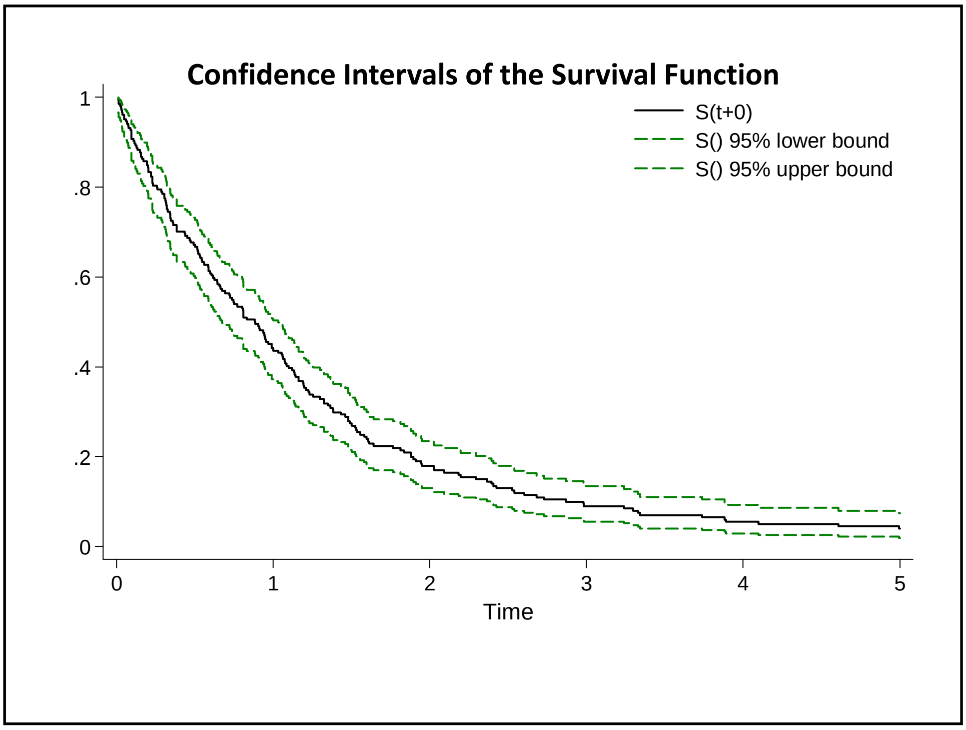
(adapted from the Stata 10 [ST] Manual p. 356)

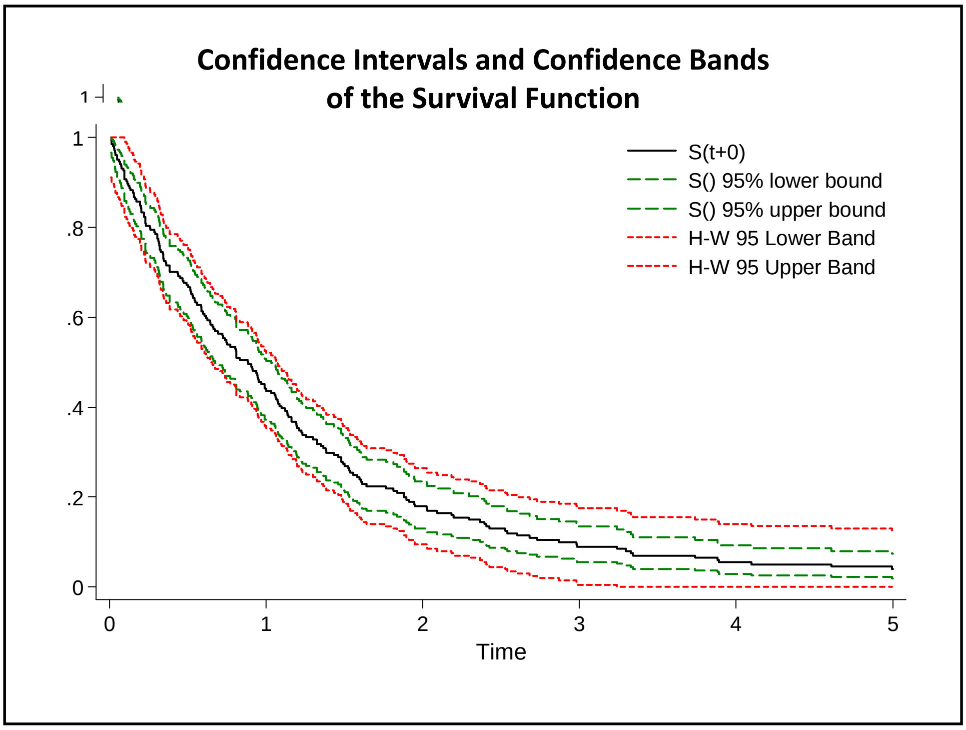
- The confidence intervals (CI) are valid at a single time point.
- A common incorrect use is to estimate CI at all time points and connect their endpoints drawing two curves. The area between the two curves is interpreted as having, for example, the 95% confidence to contain the entire survival function.
- Rather, the so-called confidence bands (not yet available within Stata) are the appropriate limits.
- A new Stata command, **-stcband-** , allows to compute these confidence bands for the survival and the cumulative hazard function.

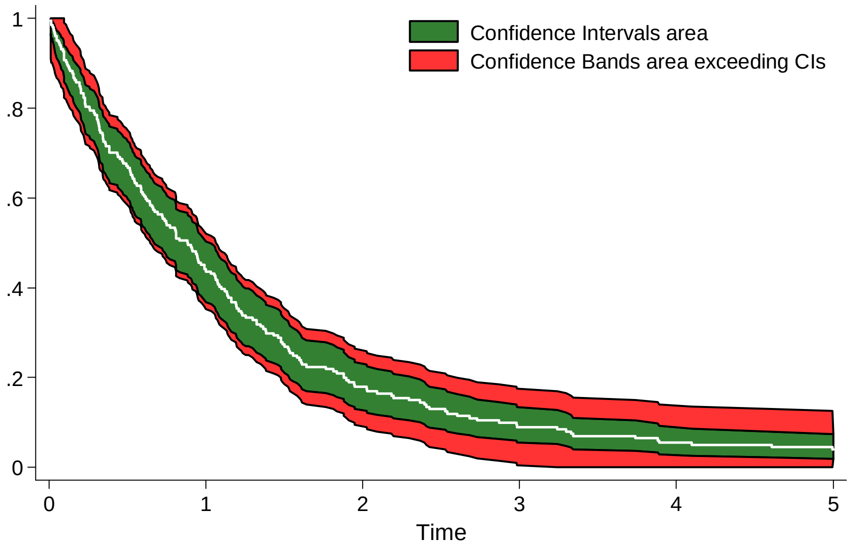
An example illustrates the difference between confidence intervals and confidence bands.

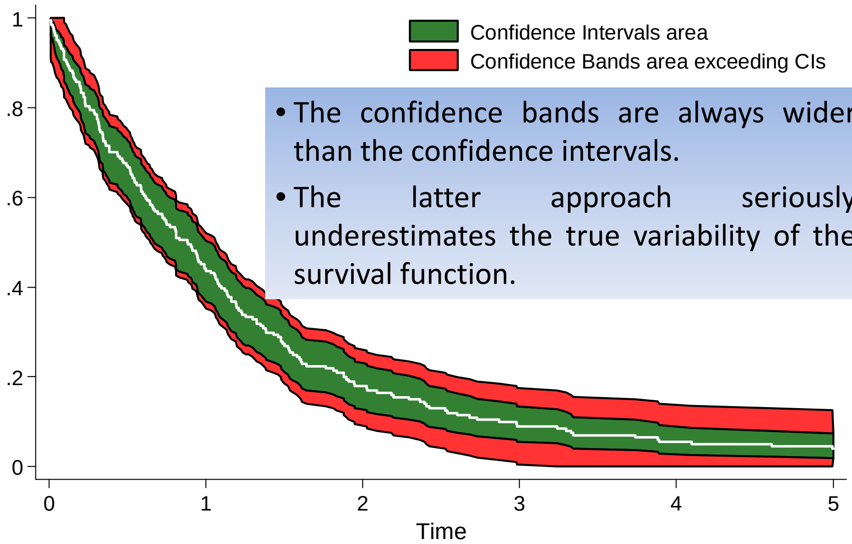
```
use rectum,clear
stset time,f(status) scale(12)

       205  total obs.
         0  exclusions
------------------------------------------------------------------
       205  obs. remaining, representing
       195  failures in single record/single failure data
  247.1417  total analysis time at risk, at risk from t = 0
                             earliest observed entry t = 0
                                  last observed exit t = 5
```

The “rectum” dataset includes 205 patients with advanced rectum cancer, of whom 195 died within 5 years (the time is in months) from the diagnosis


Confidence Intervals of the Survival Function
S(t+0)
S() 95% lower bound
S() 95% upper bound
1
.8
.6
.4
.2
0
0
1
2
3
4
5
Time


Confidence Intervals and Confidence Bands
of the Survival Function
S(t+0)
S() 95% lower bound
S() 95% upper bound
H-W 95 Lower Band
H-W 95 Upper Band
1
.8
.6
.4
.2
0
0
1
2
3
4
5
Time





- The confidence bands are always wider than the confidence intervals.
- The latter approach seriously underestimates the true variability of the survival function.

- Two methods are available to construct the confidence bands. The first has been proposed by Hall and Wellner (1980) (HW). The second, proposed by Nair (1984), is called "equal precision" (EP) [1, 2].
- To construct the confidence bands, we must use the confidence coefficients taken from special distributions.
- These coefficients are reported in the tables C.3 (Equal Precision) and C.4 (Hall and Wellner) of the Klein and Moeschberger's book[1].
- The values in the tables C.3 and C.4 have been stored in two data files: NairTables.dta and HallWellnerTables.dta

To compute confidence bands, **`-stcband-`** works as follows:

- first, four appropriate values are selected from one of these files;
- then, the selected values are linearly interpolated to determine the exact coefficient to be used.

For each method we have three possible forms of confidence bands:

- Linear
- Log-minus-log transformed (for short denoted "log")
- Arcsine square-root transformed (for short denoted "arcsine").

- Some comment about the differences and the properties of each approach is addressed at the end of the next section.

Validation of the estimates and examples

- Checks have been made to validate the new command using the “rectum” dataset
- The results obtained by -stcband- and by km.ci R function(3) were compared. As shown in the following tables, the two commands reach perfect agreement

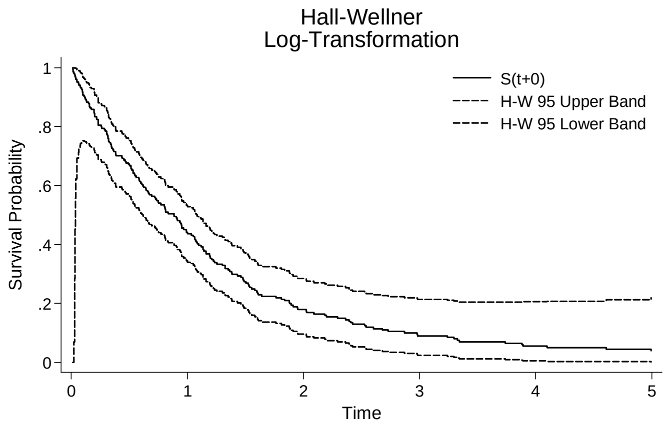
	HALL - WELLNER							
	Linear				Log-minus-log			
Time	High		Low		High		Low	
	stcband	km.ci	stcband	km.ci	stcband	km.ci	stcband	km.ci
0.108	0.9969	0.997	0.8071	0.8071	0.9635	0.964	0.7512	0.7512
0.114	0.9920	0.992	0.8022	0.8022	0.9599	0.96	0.7500	0.7500
0.122	0.9871	0.987	0.7973	0.7973	0.9561	0.956	0.7484	0.7484
0.128	0.9822	0.982	0.7924	0.7924	0.9523	0.952	0.7464	0.7464
0.133	0.9773	0.977	0.7875	0.7875	0.9484	0.948	0.7441	0.7441
2.539	0.2156	0.216	0.0232	0.0232	0.2335	0.233	0.0448	0.0448
2.603	0.2106	0.211	0.0182	0.0182	0.2297	0.23	0.0410	0.0410
2.681	0.2057	0.206	0.0132	0.0132	0.2261	0.226	0.0372	0.0372
2.731	0.2007	0.201	0.0082	0.0082	0.2226	0.223	0.0335	0.0335
2.869	0.1958	0.196	0.0032	0.0032	0.2193	0.219	0.0299	0.0299

	EQUAL PRECISION							
	Linear				Log-minus-log			
Time	High		Low		High		Low	
	stcband	km.ci	stcband	km.ci	stcband	km.ci	stcband	km.ci
0.047	0.9998	0.9998	0.9023	0.9023	0.9821	0.982	0.8700	0.8700
0.061	0.9971	0.9971	0.8952	0.8952	0.9793	0.979	0.8637	0.8637
0.067	0.9943	0.9943	0.8881	0.8881	0.9764	0.976	0.8575	0.8575
0.075	0.9915	0.9915	0.8812	0.8812	0.9735	0.973	0.8513	0.8513
0.078	0.9885	0.9885	0.8744	0.8744	0.9705	0.970	0.8451	0.8451
3.881	0.1136	0.1136	0.0058	0.0058	0.1292	0.129	0.0206	0.0206
3.892	0.1065	0.1065	0.0030	0.0030	0.1226	0.123	0.0179	0.0179
4.097	0.0992	0.0992	0.0003	0.0003	0.1160	0.116	0.0153	0.0153
4.608	0.0918	0.0918	0.0	-0.0023	0.1092	0.109	0.0128	0.0128
4.994	0.0843	0.0843	0.0	-0.0047	0.1023	0.102	0.0105	0.0105

Examples

After the -stset- statement, graphing the survival function with the confidence bands is straightforward:

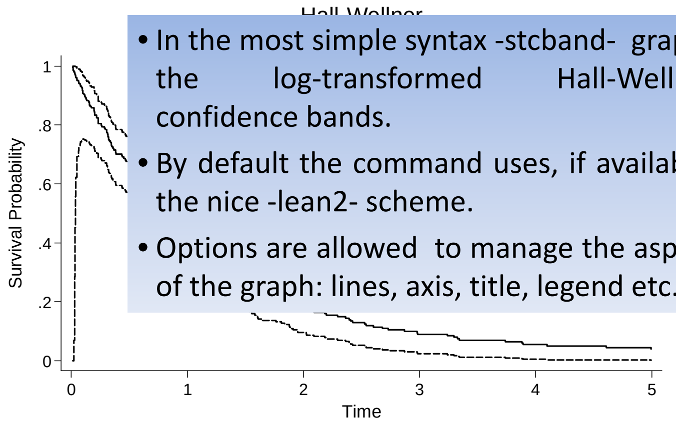
```
stcband, title("Hall-Wellner" "Log-Transformation")
```



Examples

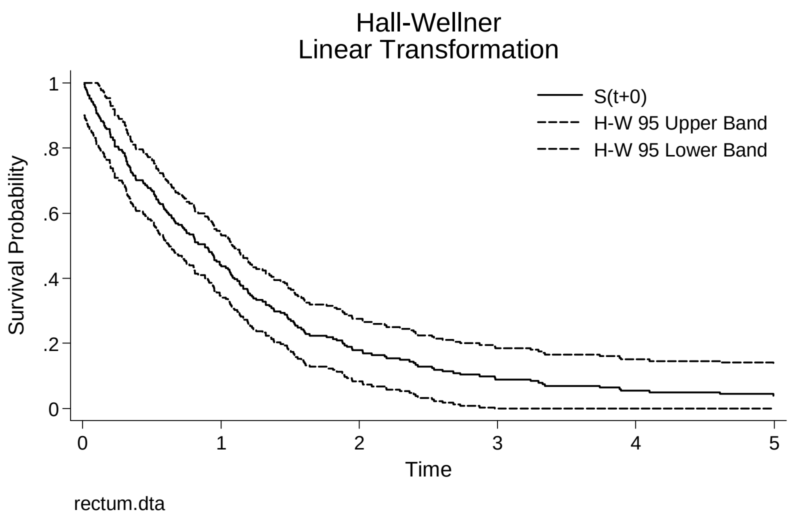
After the -stset- statement, graphing the survival function with the confidence bands is straightforward:

```
stcband, title("Hall-Wellner" "Log-Transformation")
```

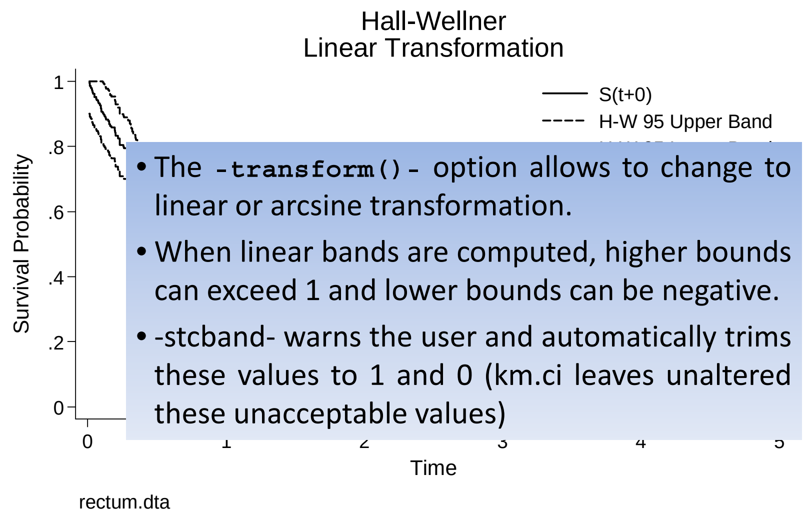


- In the most simple syntax -stcband- graphs the log-transformed Hall-Wellner confidence bands.
- By default the command uses, if available, the nice -lean2- scheme.
- Options are allowed to manage the aspect of the graph: lines, axis, title, legend etc.

```
stcband, transform(linear) caption(“rectum.dta”) ///
         title("Hall-Wellner" "Linear Transformation")
```

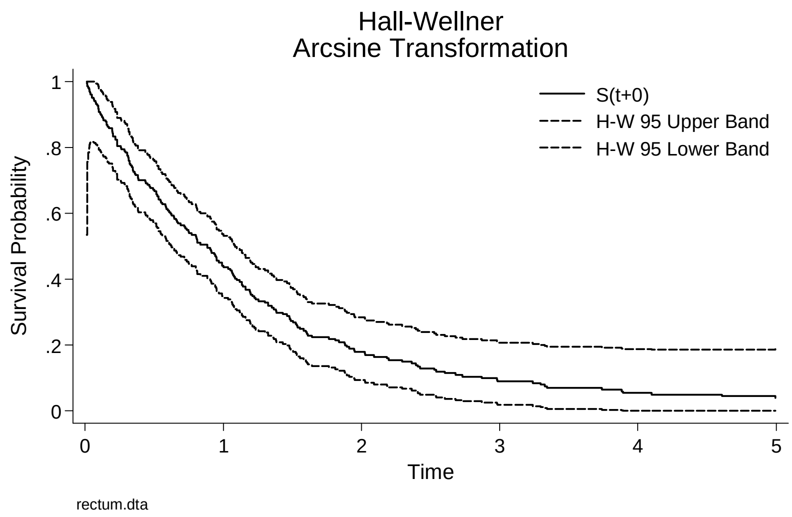


```
stcband, transform(linear) caption(“rectum.dta”) ///
         title("Hall-Wellner" "Linear Transformation")
```



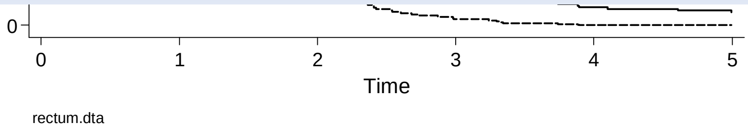
- The **-transform()-** option allows to change to linear or arcsine transformation.
- When linear bands are computed, higher bounds can exceed 1 and lower bounds can be negative.
- -stcband- warns the user and automatically trims these values to 1 and 0 (km.ci leaves unaltered these unacceptable values)

```
stcband, transform(arcsine) note(“rectum.dta”) ///
         title("Hall-Wellner" “Arcsine Transformation")
```

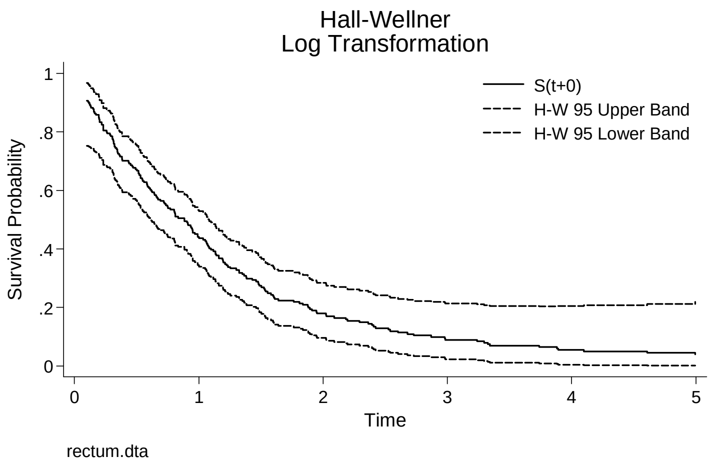


```
stcband, transform(arcsine) note(“rectum.dta”) ///
         title("Hall-Wellner" “Arcsine Transformation")
```

- Often, the arcsine and the log transformed confidence bands are both large at the beginning of the survival curve.
- This result is apparently anomalous. In this tract of the survival function, in fact, we expect the confidence bands to be shorter than in the rest of the curve.
- This happens, however, in the Hall-Wellner method alone and depends on the formulae applied.
- To circumvent this problem we can specify a lower time limit slightly greater than the minimum observed time.

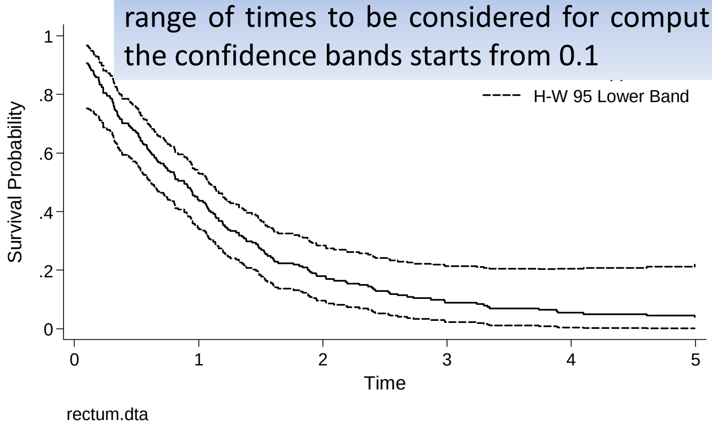



```
stcband, tlower(0.1) ......
```



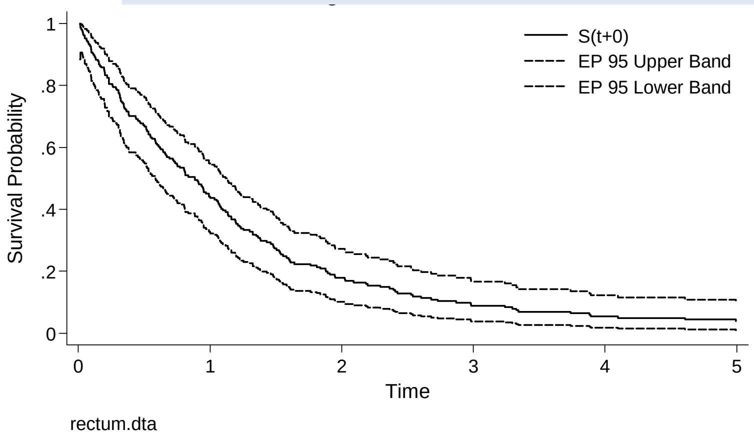
```
stcband, tlower(0.1) ……
```

In the option –tlower()- we specified that the range of times to be considered for computing the confidence bands starts from 0.1

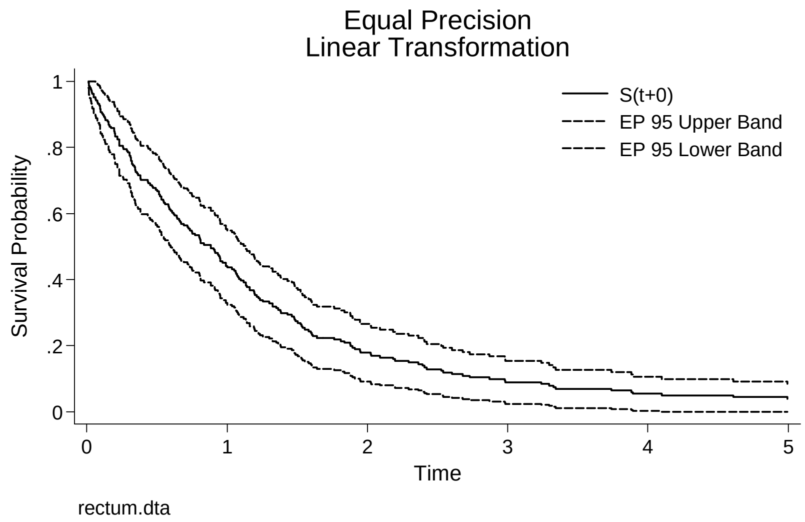



```
stcband, nair note("rectum.dta") ///
       title("Equal Precision" "Log-Transformation")
```

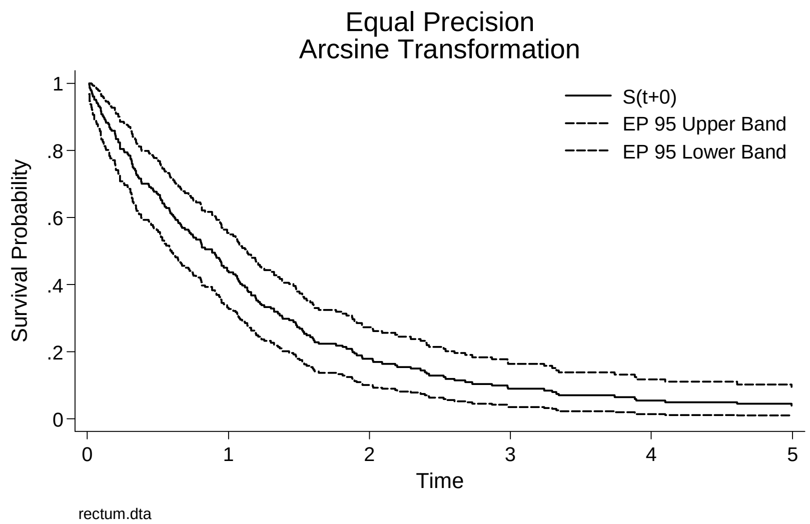
The -nair- option switches -stcband- to graph Equal Precision confidence bands



```
stcband, nair transform(linear) note("rectum.dta") ///
      title("Equal Precision" "Linear Transformation")
```




```
stcband, nair transform(arcsine) note("rectum.dta") ///
      title("Equal Precision" "Arcsine Transformation")
```



Comparing Methods and Transformations

- -stcband- can save lower and higher limits of the confidence bands by specifying the options -genhi(*newvarname*)- and -genlo(*newvarname*)-.

- After saving the estimates obtained by the Hall-Wellner and Equal Precision methods, a graph can be easily produced to compare either methods:

```
stcband, nograph genhi(HW_hi) genlo(HW_lo) tlower(0.1)
```

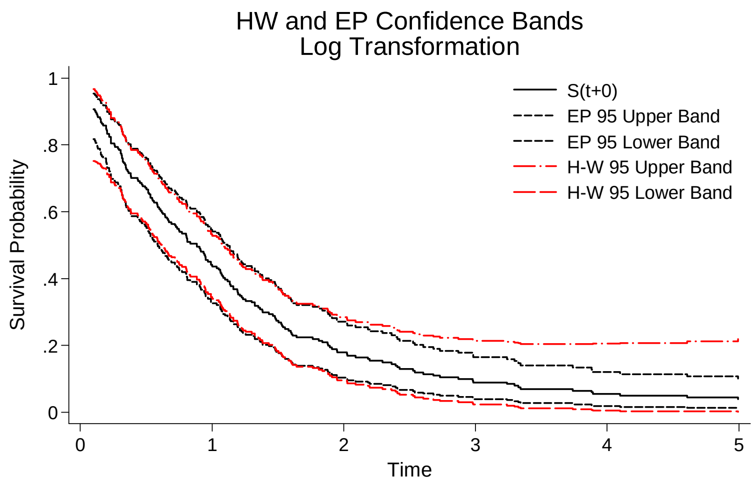
The option **-nograph-** suppresses the graph to be shown. The higher and lower limits of the log-transformed confidence bands are saved in the variables *HW_hi* and *HW_lo* .

```
stcband, nair  tlower(0.1) plot(line HW_hi HW_lo _t , ///
       sort c(J J) lc(red red)) title("HW and EP ///
       Confidence Bands" "Log Transformation")
```

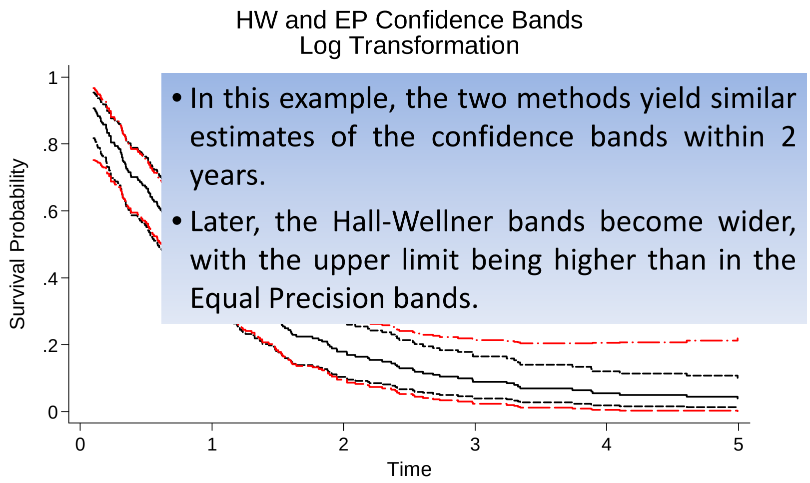
-stcband- and the **-nair-** option graph the Equal Precision confidence bands.

The **-plot()-** option overlaps the graph with the previous estimates.

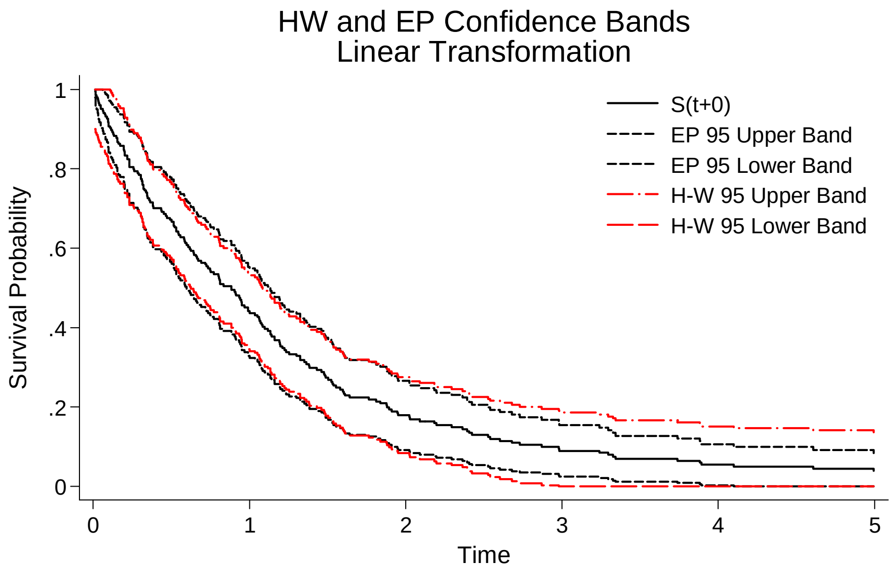
```
stcband, nair  tlower(0.1) plot(line HW_hi HW_lo _t , ///
       sort c(J J) lc(red red)) title("HW and EP ///
       Confidence Bands" "Log Transformation")
```

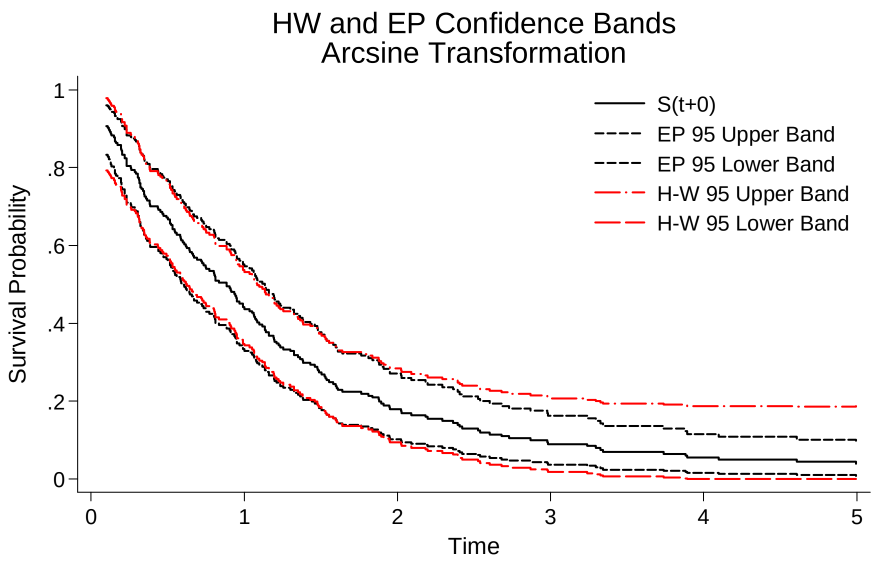


```
stcband, nair  tlower(0.1) plot(line HW_hi HW_lo _t , ///
       sort c(J J) lc(red red)) title("HW and EP ///
       Confidence Bands" "Log Transformation")
```



- In this example, the two methods yield similar estimates of the confidence bands within 2 years.
- Later, the Hall-Wellner bands become wider, with the upper limit being higher than in the Equal Precision bands.





Now, let us consider the Hall-Wellner method and compare in the same graph the linear, log and arcsine transformed confidence bands.

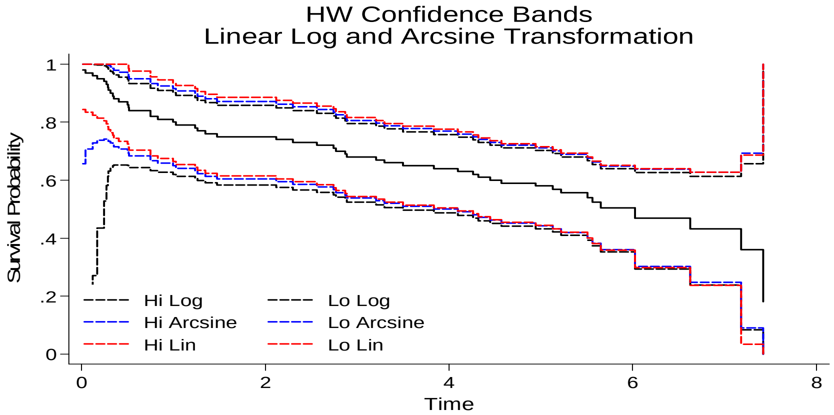
Given that the curves from the rectum data set overlap, we used the example dataset WHAS100, presented in the book Applied Survival Analysis(4):

```
use e:\whas100
stset  lenfol,f(status) scale(365.25)

stcband, nograph transform(arcsine) ///
        genhi(HiArc) genlo(LoArc)

stcband, nograph transform(linear) ///
        genhi(HiLin) genlo(LoLin)
```

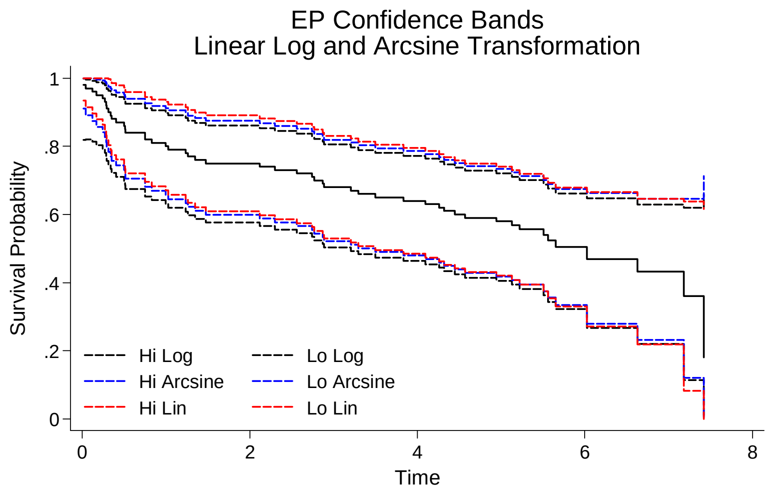
Linear and arcsine transformed estimates are saved without graphing

Now we estimates log transformed confidence bands and graph them together with the previous estimates:

```
stcband, plot(line HiArc LoArc HiLin LoLin _t , sort ///
       c(J J J J) lc(blue blue red red) lp(- - - -)) ///
       legend(label(2 "Hi Log") label(3 "Lo Log") ///
       label(4 "Hi Arcsine") label(5 "Lo Arcsine") ///
       label(6 "Hi Lin") label(7 "Lo Lin") ///
       pos(7) ring(0) rows(3) order(2 3 4 5 6 7) )
```



- At the beginning of the follow-up time, the arcsine and (more) the log transformed bands are wider than the linear ones.
- Even the linear confidence bands have a problem in this tract: the higher limit is automatically trimmed to 1 by **-stcband-**
- In the rest of the curve it is hard to see relevant differences .



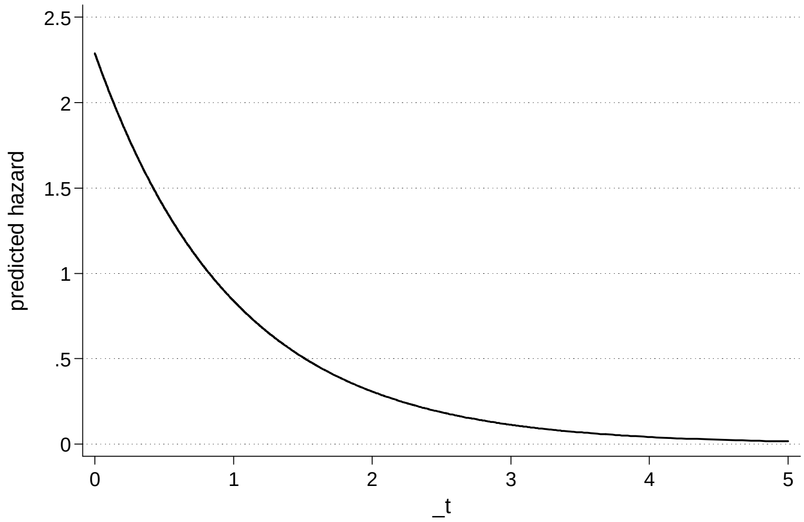
For the Equal Precision method the aforementioned problem at the start of the follow-up does not exist.

EQUAL PRECISION	HALL-WELLNER
• The confidence bands are proportional to the pointwise confidence intervals: identical formulae are applied to calculate confidence bands and intervals, but the Z coefficient in the CI formula is replaced by a different coefficient in the Equal Precision formula. • Borgan and Liestol [5] studied the coverage probabilities of the confidence bands. On this basis they recommend arcsine transformed confidence bands. The linear bounds should be avoided.	• The confidence bands are not proportional to the pointwise confidence intervals: ad hoc formulae are applied. • Anomalous values of the lower confidence band are seen at the start of the follow-up when log or arcsine transformations are used. Therefore, the initial observed times should be excluded. • Linear, log and arcsine transformed confidence bands work reasonably well with as few as 20 events [5].

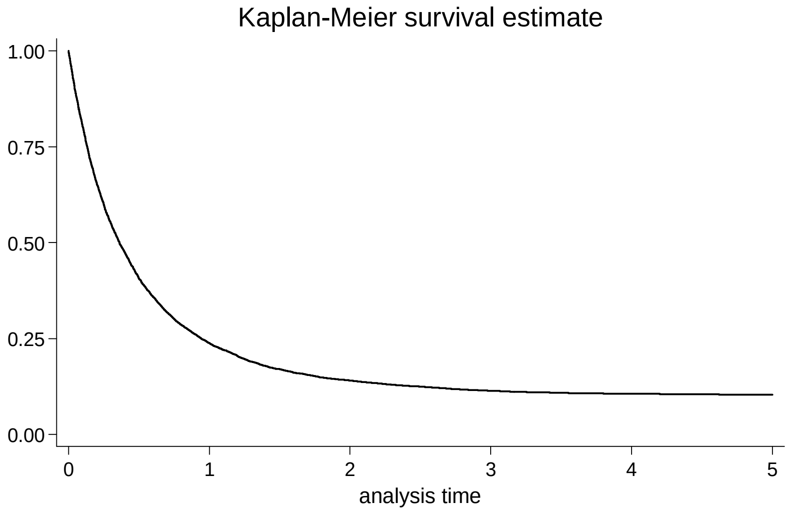
Coverage probabilities

- Using -stcband- and the -bootstrap- capabilities of Stata, a personal check of the coverage probabilities of the various approaches to estimate the confidence bands has been done.
- Briefly, two simulated data sets have been generated. The first follows a Gompertz distribution, the second a log-logistic distribution [6, 7].
- In the former distribution, the scale and shape parameters have been chosen to approximately reproduce the survival of a highly malignant tumor (lung, pancreas).
- In the latter, the scale and shape parameters mimic the survival experience of a low malignant tumor like the breast cancer.

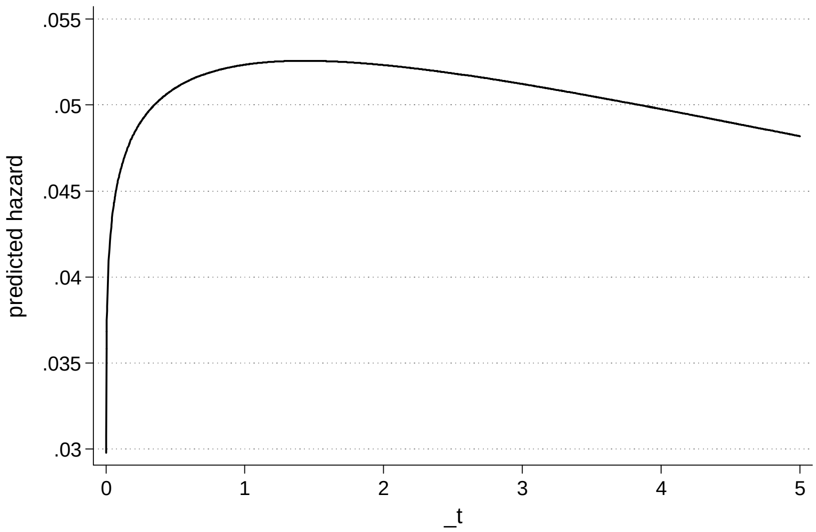
Simulated data set: Gompertz distribution



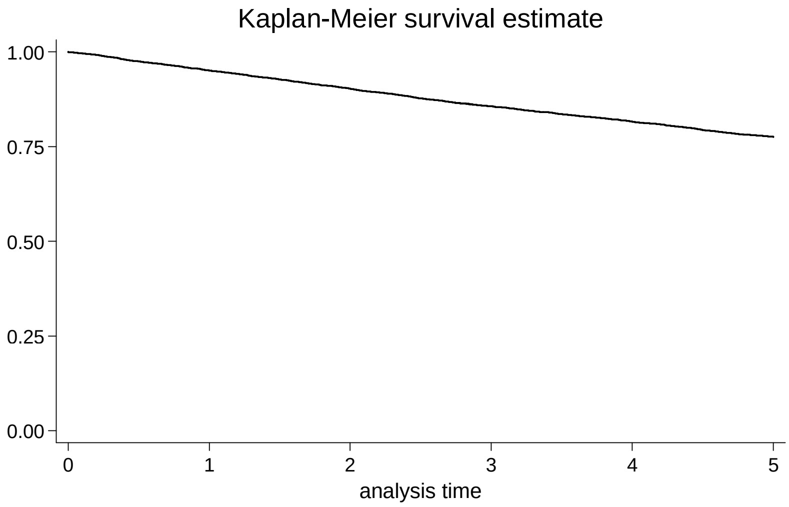
Simulated data set: Gompertz distribution



Simulated data set: Log-logistic distribution



Simulated data set: Log-logistic distribution



BOOTSTRAP

- The survival function in the simulated data has been saved in a variable. This function should represent the population (true) survival function: S_p.
- 1000 replicates has been done.
- In each sample the higher and lower limits of the confidence bands have been estimated according to 6 (2 methods x 3 transformations) different approaches.
- Then, an -assert- statement verifies whether the confidence bands encompass S_p.
- This also allows the coverage probabilities of the confidence intervals to be checked.

Each replication returns 7 results (scalars):

- r1-r6 assume value 1 if the confidence bands encompass S_p, 0 otherwise
- r7 assumes value 1 if the confidence intervals encompass S_p, 0 otherwise.

bandboot- returns :	r1=1 if EP log	bands encompass the survival function
	r2=1 if EP arcsine	" " " "
	r3=1 if EP linear	" " " "
	r4=1 if H-W log	" " " "
	r5=1 if H-W arcsine	" " " "
	r6=1 if H-W linear	" " " "
	r7=1 if Pointwise	" " " "

RESULTS - GOMPERTZ DISTRIBUTION

	Obs	COVERAGE PROBABILITIES
Eq Prec LOG-LOG	10000	**0.935**
Eq Prec ARCSINE	10000	**0.946**
Eq Prec LINEAR	10000	**0.918**
HALL-WELLNER LOG-LOG	10000	**0.951**
HALL-WELLNER ARCSINE	10000	**0.952**
HALL-WELLNER LINEAR	10000	**0.951**
POINTWISE CONFIDENCE INTERVALS	10000	**0.297**

RESULTS - GOMPERTZ DISTRIBUTION

	Obs	COVERAGE PROBABILITIES
Eq Prec		
Eq Prec		
Eq Prec		
HALL-WE		
HALL-WE		
HALL-WE		
POINTWI INTERVALS	10000	**0.297**

- The Equal Precision linear confidence bands performs slightly worse than the other approaches.
- The coverage probabilities of the Hall-Wellner method correspond exactly to the nominal value without differences among linear, log or arcsine transformed form.
- The pointwise confidence intervals strongly underestimate the variability of the survival function.

RESULTS – LOG-LOGISTIC DISTRIBUTION

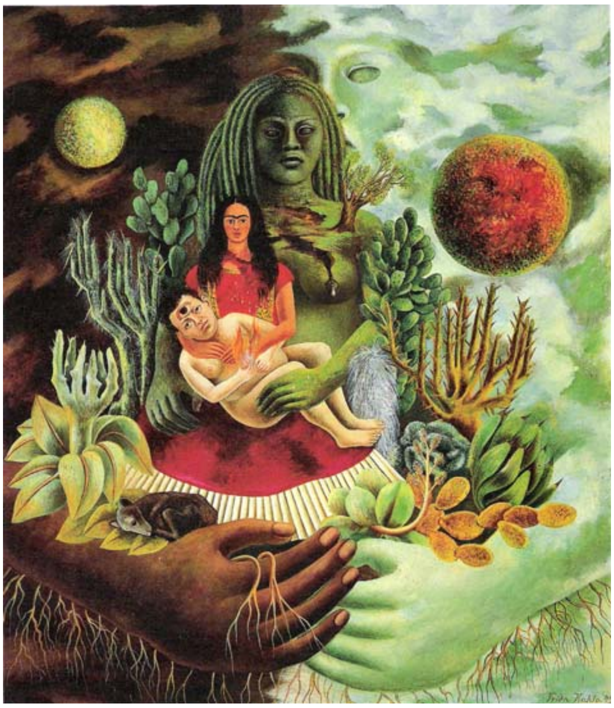
	Obs	COVERAGE PROBABILITIES
Eq Prec LOG-LOG	10000	**0.915**
Eq Prec ARCSINE	10000	0.926
Eq Prec LINEAR	10000	**0.887**
HALL-WELLNER LOG-LOG	10000	**0.949**
HALL-WELLNER ARCSINE	10000	**0.949**
HALL-WELLNER LINEAR	10000	**0.949**
POINTWISE CONFIDENCE INTERVALS	10000	**0.457**

RESULTS – LOG-LOGISTIC DISTRIBUTION

	Obs	COVERAGE PROBABILITIES
Eq Prec		
Eq Prec		
Eq Prec		
HALL-WEL		
HALL-WEL		
HALL-WEL		
POINTWIS		

In this different context the results look corresponding to the previous one:

- the coverage probabilities of the Equal Precision linear confidence bands performs worse
- the Hall-Wellner method yields better results than the Equal Precision
- the coverage probabilities of the pointwise confidence intervals are again unsatisfactory at all.



CONCLUSIONS

- In clinical and epidemiological settings, the confidence bands should be used when dealing with the variability of the survival function.
- To this aim the confidence intervals are inappropriate, as confirmed by our results of two simulations. Their use is no longer justified by the unavailability of software estimating the appropriate confidence bands.
- The new Stata command -stcband- makes available the estimates of the confidence bands of the survival function according to 2 methods and 3 transformations.
- Although not illustrated in this talk, the -na- option [1, 8] of -stcband- allows confidence bands for the cumulative hazard function to be estimated too.

The full syntax of **-stcband-** is as follows:

```
stcband [if] [in] [,
      nair tower(#) tupper(#) na
      transform(linear log arcsine)
      genlow(newvar) genhigh(newvar) level(#->90-95-99)
      nograph twoway_options   ]
```

- The new command is also provided with a help file in which the user can run an example, taken from Klein and Moeschberger's book(1), by clicking on the viewer window.
- **-stcband-** is available for download from the SSC-Archive.

References

1. Klein J.P. and Moeschberger M.L. Survival Analysis: techniques for Censored and Truncated Data (2nd ed.), pp. 104-117. New York: Springer-Verlag, 2003.
2. Borgan O. The Kaplan-Meier estimator in Encyclopedia of Biostatistics (eds. P. Armitage and T. Colton), vol 3, pp. 2154-60. Chichester: Wiley, 1998
3. Strobl R. The km.ci package. Version 0.5-1, 2007. www.mirrorservice.org/sites/lib.stat.cmu.edu/R/CRAN/doc/packages/km.ci.pdf
4. Hosmer D.W., Lemeshow S. and May S. Applied Survival Analysis (2nd ed.), pp. 27-35. Hoboken, New Jersey: John Wiley & Sons, 2008.
5. Borgan O. and Liestol K. A note on confidence intervals and bands for the survival function based on transformations. Scand. J. Statist. 17: 35-41, 1990.
6. Bender R., Aug… Cox proportiona…
7. Burton A. Altm… studies in medi…
8. Borgan O. The Nelson-Aalen estimator in Encyclopedia of Biostatistics (eds. P. Armitage and T. Colton), vol 4, pp. 2967-72. Wiley, Chichester, 1998.

I wish to thank Maarten Buis for his brilliant advices in constructing the simulations checking the coverage probabilities of the confidence bands.

References

1. Klein J.P. and Moeschberger M.L. Survival Analysis: techniques for Censored and Truncated Data (2nd ed.), pp. 104-117. New York: Springer-Verlag, 2003.
2. Borgan O. The Kaplan-Meier estimator in Encyclopedia of Biostatistics (eds. P. Armitage and T. Colton), vol 3, pp. 2154-60. Chichester: Wiley, 1998
3. Strobl R. The km.ci package. Version 0.5-1, 2007. www.mirrorservice.org/sites/lib.stat.cmu.edu/R/CRAN/doc/packages/km.ci.pdf
4. Hosmer D.W., Lemeshow S. and May S. Applied Survival Analysis (2nd ed.), pp. 27-35. Hoboken, New Jersey: John Wiley & Sons, 2008.
5. Borgan O. and Liestol K. A note on confidence intervals and bands for the survival function based on transformations. Scand. J. Statist. 17: 35-41, 1990.
6. Bender R., Augustin T. and Blettner M. Generating survival times to simulate Cox proportional hazards model. Statist. Med. 2005; 24: 1713-1723
7. Burton A. Altman D.G., Royston P. and Holder R.L. The design of simulations studies in medical statistics. Statist. Med. 2006: 25: 42279-4292.
8. Borgan O. The Nelson-Aalen estimator in Encyclopedia of Biostatistics (eds. P. Armitage and T. Colton), vol 4, pp. 2967-72. Wiley, Chichester, 1998.

Paintings from Frida Kahlo

Thanks