Title

stcurve — Plot survivor, hazard, cumulative hazard, or cumulative incidence function

Syntax   Menu   Description   Options
[Remarks and examples](Remarks-and-examples)   References   [Also see](Also-see)

Syntax

    stcurve [, *options*]

options	Description
Main	
* <u>surv</u>ival	plot survivor function
* <u>hazard</u>	plot hazard function
* <u>cumhaz</u>	plot cumulative hazard function
* <u>cif</u>	plot cumulative incidence function
at(*varname*=# [*varname*=#...])	value of the specified covariates and mean of
[at1(*varname*=# [*varname*=#...])	unspecified covariates
[at2(*varname*=# [*varname*=#...])	
[...]]]	
Options	
<u>alpha1</u>	conditional frailty model
<u>uncond</u>itional	unconditional frailty model
<u>range</u>(# #)	range of analysis time
<u>outf</u>ile(*filename* [, replace])	save values used to plot the curves
<u>width</u>(#)	override "optimal" width; use with hazard
<u>kernel</u>(*kernel*)	kernel function; use with hazard
<u>noboundary</u>	no boundary correction; use with hazard
Plot	
connect_options	affect rendition of plotted survivor, hazard, or cumulative hazard function
Add plots	
<u>addplot</u>(*plot*)	add other plots to the generated graph
Y axis, X axis, Titles, Legend, Overall	
twoway_options	any options other than by() documented in [G-3] *twoway_options*

  *One of survival, hazard, cumhaz, or cif must be specified.

    survival and hazard are not allowed after estimation with stcrreg; see [ST] **stcrreg**.

    cif is allowed only after estimation with stcrreg; see [ST] **stcrreg**.

Menu

Statistics > Survival analysis > Regression models > Plot survivor, hazard, cumulative hazard, or cumulative incidence function

Description

stcurve plots the survivor, hazard, or cumulative hazard function after stcox or streg and plots the cumulative subhazard or cumulative incidence function (CIF) after stcrreg.

Options

    ┌─── Main ───

survival requests that the survivor function be plotted. survival is not allowed after estimation with stcrreg.

hazard requests that the hazard function be plotted. hazard is not allowed after estimation with stcrreg.

cumhaz requests that the cumulative hazard function be plotted when used after stcox or streg and requests that the cumulative subhazard function be plotted when used after stcrreg.

cif requests that the cumulative incidence function be plotted. This option is available only after estimation with stcrreg.

at(*varname*=# ...) requests that the covariates specified by *varname* be set to #. By default, stcurve evaluates the function by setting each covariate to its mean value. This option causes the function to be evaluated at the value of the covariates listed in at() and at the mean of all unlisted covariates.

    at1(*varname*=# ...), at2(*varname*=# ...), ..., at10(*varname*=# ...) specify that multiple curves (up to 10) be plotted on the same graph. at1(), at2(), ..., at10() work like the at() option. They request that the function be evaluated at the value of the covariates specified and at the mean of all unlisted covariates. at1() specifies the values of the covariates for the first curve, at2() specifies the values of the covariates for the second curve, and so on.

    ┌─── Options ───

alpha1, when used after fitting a frailty model, plots curves that are conditional on a frailty value of one. This is the default for shared-frailty models.

unconditional, when used after fitting a frailty model, plots curves that are unconditional on the frailty; that is, the curve is "averaged" over the frailty distribution. This is the default for unshared-frailty models.

range(# #) specifies the range of the time axis to be plotted. If this option is not specified, stcurve plots the desired curve on an interval expanding from the earliest to the latest time in the data.

outfile(*filename* [, replace]) saves in *filename*.dta the values used to plot the curve(s).

width(#) is for use with hazard (and applies only after stcox) and is used to specify the bandwidth to be used in the kernel smooth used to plot the estimated hazard function. If left unspecified, a default bandwidth is used, as described in [R] kdensity.

kernel(*kernel*) is for use with hazard and is for use only after stcox because, for Cox regression, an estimate of the hazard function is obtained by smoothing the estimated *hazard contributions*. kernel() specifies the kernel function for use in calculating the weighted kernel-density estimate required to produce a smoothed hazard-function estimator. The default is kernel(Epanechnikov), yet *kernel* may be any of the kernels supported by kdensity; see [R] **kdensity**.

noboundary is for use with hazard and applies only to the plotting of smoothed hazard functions after stcox. It specifies that no boundary-bias adjustments are to be made when calculating the smoothed hazard-function estimator. By default, the smoothed hazards are adjusted near the boundaries; see [ST] **sts graph**. If the epan2, biweight, or rectangular kernel is used, the bias correction near the boundary is performed using boundary kernels. For other kernels, the plotted range of the smoothed hazard function is restricted to be inside of one bandwidth from each endpoint. For these other kernels, specifying noboundary merely removes this range restriction.

⎣ Plot ⎤

connect_options affect the rendition of the plotted survivor, hazard, or cumulative hazard function; see [G-3] **connect_options**.

⎣ Add plots ⎤

addplot(*plot*) provides a way to add other plots to the generated graph; see [G-3] **addplot_option**.

⎣ Y axis, X axis, Titles, Legend, Overall ⎤

twoway_options are any of the options documented in [G-3] **twoway_options**, excluding by(). These include options for titling the graph (see [G-3] **title_options**) and for saving the graph to disk (see [G-3] **saving_option**).

Remarks and examples   stata.com
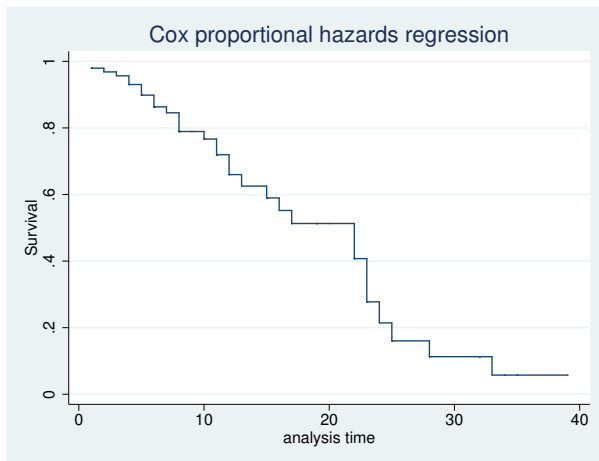
Remarks are presented under the following headings:

stcurve after stcox

After fitting a Cox model, stcurve can be used to plot the estimated hazard, cumulative hazard, and survivor functions.
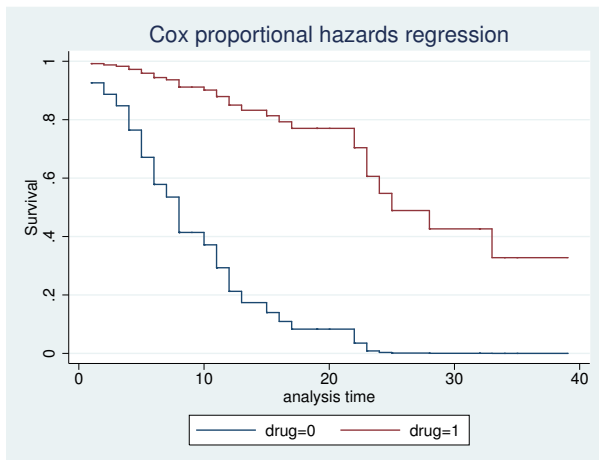
▷ Example 1

```
. use http://www.stata-press.com/data/r13/drugtr
(Patient Survival in Drug Trial)

. stcox age drug
  (output omitted)

. stcurve, survival
```


By default, the curve is evaluated at the mean values of all the predictors, but we can specify other values if we wish.

```
. stcurve, survival at1(drug=0) at2(drug=1)
```


In this example, we asked for two plots, one for the placebo group and one for the treatment group. For both groups, the value of `age` was held at its mean value for the overall estimation sample.

See Cefalu (2011) for a Stata command to plot the survivor or cumulative hazard function with pointwise confidence intervals.

◁

▷ Example 2

stcurve can also be used to plot estimated hazard functions. The hazard function is estimated by a kernel smooth of the estimated hazard contributions; see [ST] **sts graph** for details. We can thus customize the smooth as we would any other; see [R] **kdensity** for details.

```
. stcurve, hazard at1(drug=0) at2(drug=1) kernel(gauss) yscale(log)
```
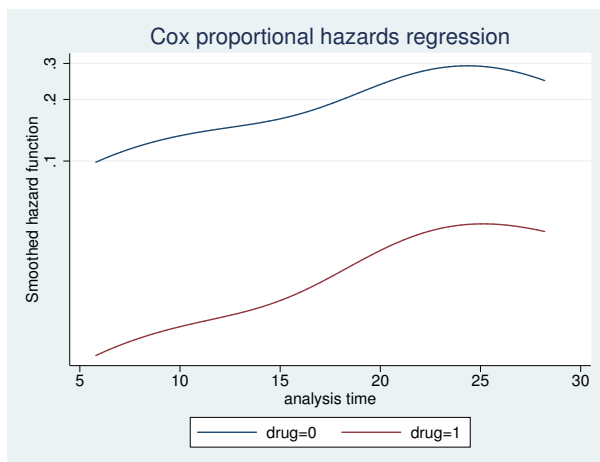

For the hazard plot, we plotted on a log scale to demonstrate the proportionality of hazards under this model; see the technical note below on smoothed hazards.

◁

❑ Technical note

For survivor or cumulative hazard estimation, stcurve works by first estimating the baseline function and then modifying it to adhere to the specified (or by default, mean) covariate patterns. As mentioned previously, *baseline* (when all covariates are equal to zero) must correspond to something that is meaningful and preferably in the range of your data. Otherwise, stcurve could encounter numerical difficulties. We ignored our own advice above and left age unchanged. Had we encountered numerical problems, or funny-looking graphs, we would have known to try shifting age so that age==0 was in the range of our data.

For hazard estimation, stcurve works by first transforming the estimated hazard contributions to adhere to the necessary covariate pattern and then applying the smooth. When you plot multiple curves, each is smoothed independently, although the same bandwidth is used for each.

The smoothing takes place in the hazard scale and not in the log hazard-scale. As a result, the resulting curves will look nearly, but not exactly, parallel when plotted on a log scale. This inexactitude is a product of the smoothing and should not be interpreted as a deviation from the proportional-hazards assumption; stcurve (after stcox) assumes proportionality of hazards and will reflect this in the produced plots. If smoothing were a perfect science, the curves would be parallel when plotted on a log scale. If you encounter estimated hazards exhibiting severe disproportionality, this may signal a numerical problem as described above. Try recentering your covariates so that baseline is more reasonable.

❑

stcurve after streg

stcurve is used after streg to plot the fitted survivor, hazard, and cumulative hazard functions. By default, stcurve computes the means of the covariates and evaluates the fitted model at each time in the data, censored or uncensored. The resulting plot is therefore the survival experience of a subject with a covariate pattern equal to the average covariate pattern in the study. You can produce the plot at other values of the covariates by using the at() option or specify a time range by using the range() option.

▷ Example 3

We pick up where example 6 of [ST] **streg** left off. The cancer dataset we are using has three values for variable drug: 1 corresponds to placebo, and 2 and 3 correspond to two alternative treatments. Using the cancer data with drug remapped to form an indicator of treatment, let's fit a loglogistic regression model and plot its survival curves. We can perform a loglogistic regression by issuing the following commands:

```
. use http://www.stata-press.com/data/r13/cancer
(Patient Survival in Drug Trial)
. replace drug = drug==2 | drug==3          // 0, placebo : 1, nonplacebo
(48 real changes made)
. stset studytime, failure(died)
 (output omitted )
. streg age drug, distribution(llogistic) nolog

        failure _d:  died
  analysis time _t:  studytime

Loglogistic regression -- accelerated failure-time form

No. of subjects =            48                    Number of obs   =         48
No. of failures =            31
Time at risk    =           744
                                                   LR chi2(2)      =      35.14
Log likelihood  =    -43.21698                     Prob > chi2     =     0.0000
```

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
age	-.0803289	.0221598	-3.62	0.000	-.1237614	-.0368964
drug	1.420237	.2502148	5.68	0.000	.9298251	1.910649
_cons	6.446711	1.231914	5.23	0.000	4.032204	8.861218
/ln_gam	-.8456552	.1479337	-5.72	0.000	-1.1356	-.5557105
gamma	.429276	.0635044			.3212293	.5736646

Now we wish to plot the survivor and the hazard functions:
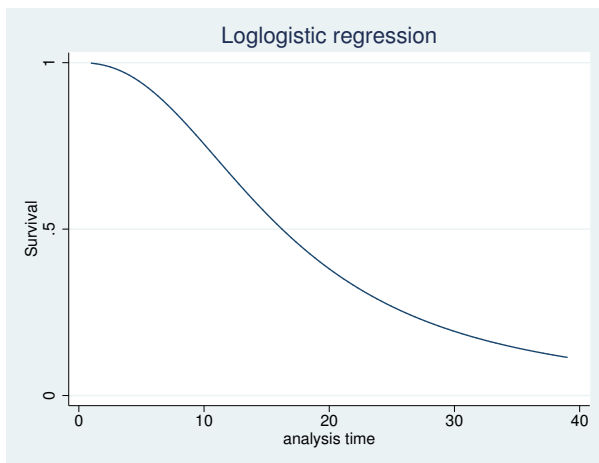
```
. stcurve, survival ylabels(0 .5 1)
```

Figure 3. Loglogistic survival distribution at mean value of all covariates
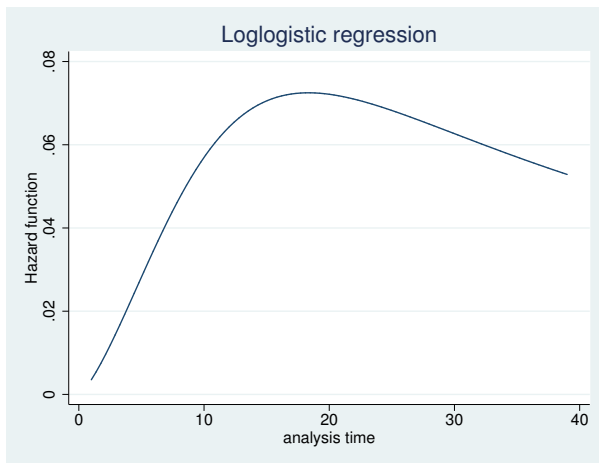
```
. stcurve, hazard
```

Figure 4. Loglogistic hazard distribution at mean value of all covariates

These plots show the fitted survivor and hazard functions evaluated for a cancer patient of average age receiving the average drug. Of course, the "average drug" has no meaning here because `drug` is an indicator variable. It makes more sense to plot the curves at a fixed value (level) of the drug. We can do this with the `at` option. For example, we may want to compare the average-age patient's survival curve under placebo (`drug==0`) and under treatment (`drug==1`).

We can plot both curves on the same graph:

```
. stcurve, surv at1(drug = 0) at2(drug = 1) ylabels(0 .5 1)
```
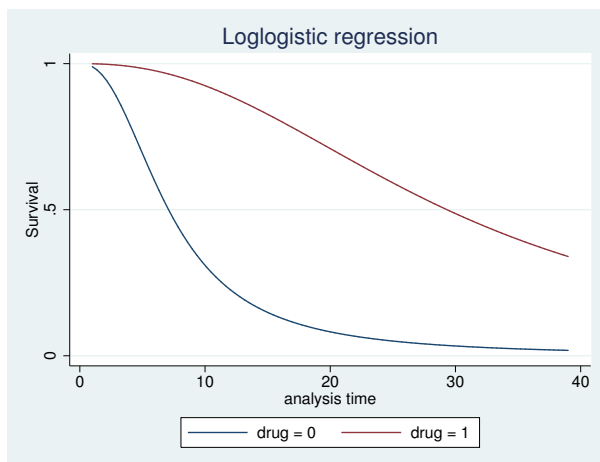

Figure 5. Loglogistic survival distribution at mean age for placebo

In the plot, we can see from the loglogistic model that the survival experience of an average-age patient receiving the placebo is worse than the survival experience of that same patient receiving treatment. We can also see the accelerated-failure-time feature of the loglogistic model. The survivor function for treatment is a time-decelerated (stretched-out) version of the survivor function for placebo.

◁

▷ Example 4

In our discussion of frailty models in [ST] **streg**, we emphasize the distinction between the individual hazard (or survivor) function and the hazard (survivor) function for the population. When significant frailty is present, the population hazard will tend to begin falling past a certain point, regardless of the shape of the individual hazard. This is due to the frailty effect—as time passes, the frailer individuals will fail, leaving a more homogeneous population comprising only the most robust individuals.

The frailty effect may be demonstrated using `stcurve` to plot the estimated hazard (both individual and population) after fitting a frailty model. Use the `alpha1` option to specify the individual hazard ($\alpha = 1$) and the `unconditional` option to specify the population hazard. Applying this to the Weibull/inverse-Gaussian shared-frailty model on the kidney data of example 11 of [ST] **streg**,

```
. use http://www.stata-press.com/data/r13/catheter, clear
(Kidney data, McGilchrist and Aisbett, Biometrics, 1991)
. stset time infect
 (output omitted )
. quietly streg age female, d(weibull) frailty(invgauss) shared(patient)
```

```
. stcurve, hazard at(female = 1) alpha1
```
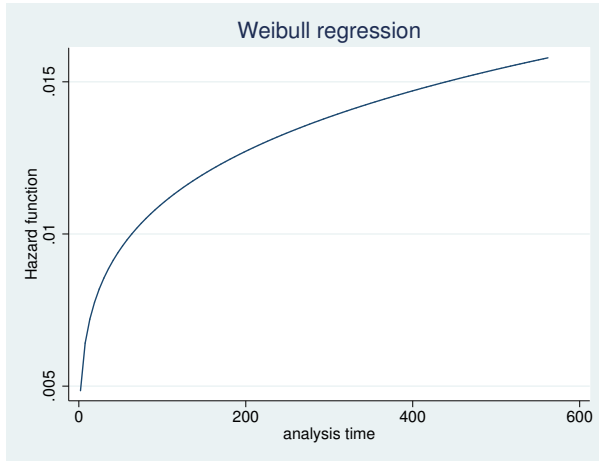

Figure 6. Individual hazard for females at mean age

Compare with

```
. stcurve, hazard at(female = 1) unconditional
```
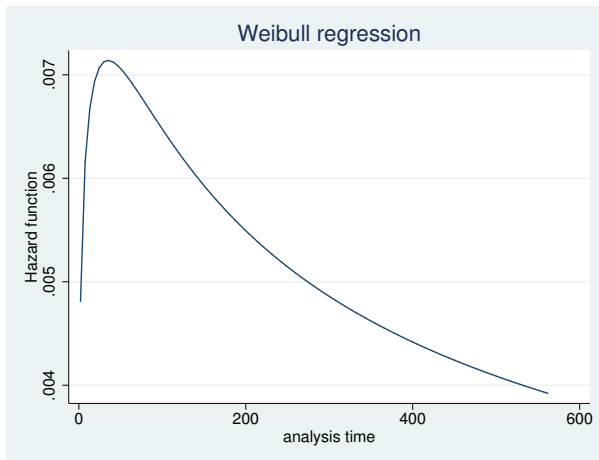

Figure 7. Population hazard for females at mean age

◁

stcurve after stcrreg

▷ Example 5

In [ST] **stcrreg**, we analyzed data from 109 patients with primary cervical cancer, treated at a cancer center between 1994 and 2000. We fit a competing-risks regression model where local relapse was the failure event of interest (`failtype == 1`), distant relapse with no local relapse was the competing risk event (`failtype == 2`), and we were interested primarily in the effect of interstitial fluid pressure (`ifp`) while controlling for tumor size and pelvic node involvement.

After fitting the competing-risks regression model, we can use `stcurve` to plot the estimated cumulative incidence of local relapses in the presence of the competing risk. We wish to compare the cumulative incidence curves for `ifp == 5` versus `ifp == 20`, assuming positive pelvic node involvement (`pelnode == 0`) and a tumor size that is the average over the data.

```
. use http://www.stata-press.com/data/r13/hypoxia
(Hypoxia study)
. stset dftime, fail(failtype==1)
  (output omitted)
. stcrreg ifp tumsize pelnode, compete(failtype==2)
  (output omitted)
. stcurve, cif at1(ifp=5 pelnode=0) at2(ifp=20 pelnode=0)
```
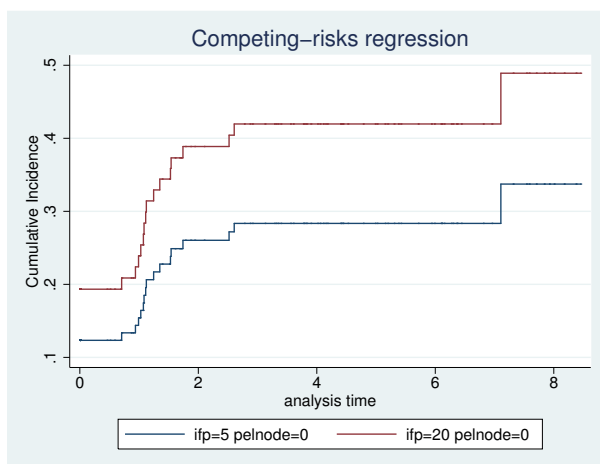

Figure 8. Comparative cumulative incidence functions

◁

References

Cefalu, M. S. 2011. Pointwise confidence intervals for the covariate-adjusted survivor function in the Cox model. *Stata Journal* 11: 64–81.

Cleves, M. A. 2000. stata54: Multiple curves plotted with stcurv command. *Stata Technical Bulletin* 54: 2–4. Reprinted in *Stata Technical Bulletin Reprints*, vol. 9, pp. 7–10. College Station, TX: Stata Press.

Also see

[ST] **stcox** — Cox proportional hazards model

[ST] **stcox postestimation** — Postestimation tools for stcox

[ST] **stcrreg** — Competing-risks regression

[ST] **stcrreg postestimation** — Postestimation tools for stcrreg

[ST] **streg** — Parametric survival models

[ST] **streg postestimation** — Postestimation tools for streg

[ST] **sts** — Generate, graph, list, and test the survivor and cumulative hazard functions

[ST] **stset** — Declare data to be survival-time data