

## Title

**intro** — Introduction to survival analysis manual

## Description

This entry describes this manual and what has changed since Stata 9. See the next entry, [ST] **survival analysis**, for an introduction to Stata's survival analysis capabilities.

## Remarks

This manual documents commands for survival analysis and epidemiological tables and is referred to as [ST] in cross-references. Following this entry, [ST] **survival analysis** provides an overview of the commands.

This manual is arranged alphabetically. If you are new to Stata's survival analysis and epidemiological tables commands, we recommend that you read the following sections first:

[ST] <b>survival analysis</b>	Introduction to survival analysis & epidemiological tables commands
[ST] <b>st</b>	Survival-time data
[ST] <b>stset</b>	Set variables for survival data

Stata is continually being updated, and Stata users are always writing new commands. To find out about the latest survival analysis features, type `search survival` after installing the latest official updates; see [R] **update**. To find out about the latest epidemiological features, type `search epi`.

## What's new

This section is intended for previous Stata users. If you are new to Stata, you may as well skip it.

- Existing estimation commands `stcox` and `streg` may now be used with the `svy:` prefix and so can fit models for complex survey data; see [ST] **stcox** and [ST] **streg**.
- New command `stpower` provides sample-size and power calculations for survival studies that use Cox proportional hazards regressions, log-rank tests for two groups, or differences in exponentially distributed hazards or log hazards.
  - `stpower cox` estimates required sample size (given power) or power (given sample size) or the minimal detectable coefficient (given power and sample size) for models with multiple covariates. The command provides options to account for possible correlation between the covariate of interest and other predictors and for withdrawal of subjects from the study. See [ST] **stpower cox**.
  - `stpower logrank` estimates required sample size (given power) or power (given sample size) or the minimal detectable hazard ratio (given power and sample size) for studies comparing survivor functions of two groups by using the log-rank test. Both the Freedman (1982) and the Schoenfeld (1981) methods are provided. The command allows for unequal allocation of subjects between the groups and possible withdrawal of subjects. Estimates can be adjusted for uniform accrual. See [ST] **stpower logrank**.

- c. `stpower exponential` estimates sample size (given power) or power (given sample size) of tests of the difference between hazards or log hazards of two groups under the assumption of exponential survivor functions (also known as the exponential test). Both the Lachin–Foulkes (1986) and Rubinstein–Gail–Santner (1981) methods are provided. Unequal group allocation, uniform or truncated exponential accrual, and different exponential losses due to follow-up in each group are allowed. See [ST] **stpower exponential**.

The `stpower` commands allow automated production of customizable tables and have options to assist with creating graphs of power curves. See [ST] **stpower**.

3. Concerning existing command `sts graph`,
  - a. New option `risktable()` places a subjects-at-risk table underneath and aligned with the survivor or hazard plot.
  - b. New option `ci` replaces old options `gwood`, `cna`, and `cihazard`. `sts graph` will choose the appropriate confidence interval on the basis of the function being graphed.
  - c. Confidence intervals are now graphed using shaded areas and new options `plotopts()` and `ciopts()` allow you to control how plots and confidence intervals look.
  - d. Overlaid confidence intervals are now allowed and are produced when new option `ci` is combined with existing option `by(varlist)`.
  - e. New option `censopts()` controls the appearance of ticks and markers produced by existing option `censored()`.
  - f. Boundary computations for smoothing hazards have been improved. New option `noboundary` specifies that no boundary correction be done.
  - g. The lower bound of the range to plot the hazard function now extends to zero.
  - h. Option `na` has been renamed `cumhaz`. `na` may still be used.

See [ST] **sts graph**. Setting `version` to less than 10 restores previous behavior.

4. For `sts list`, option `na` has been renamed `cumhaz`. `na` may be used as a synonym for `cumhaz`. See [ST] **sts list**.
5. Improvements to `stcurve` analogous to those of `sts graph` have been made.
  - a. Boundary computations for smoothing hazards have been improved. New option `noboundary` specifies that no boundary correction be done.
  - b. The lower bound of the range to plot the hazard function now extends to zero.

See [ST] **stcurve**.

6. All `st` estimation commands accept option `vce(vctype)`. As mentioned in the [U] **1.3.3 What's new in statistics (general)**, `vce(robust)` and `vce(cluster varname)` are the right ways to specify the old `robust` and `cluster()` options, and option `vce()` now allows other VCE calculations as well.
7. Existing command `predict` after `stcox` has a new option, `scores`, that allows generating variables with the partial efficient score residuals; see [ST] **stcox postestimation**.
8. Existing command `ltable` has new options `byopts()`, `plotopts()`, `plot#opts()`, and `ci#opts()` that allow for more customization of the graph. New option `ci` adds confidence intervals to the graph. See [ST] **ltable**.
9. Existing command `stphplot` has a new option `plot#opts()` that allows for more customization of the graph. See [ST] **stcox diagnostics**.

10. Existing command `stcoxkm` has new options `byopts()`, `obsopts()`, `obs#opts()`, `predopts()`, and `pred#opts()` that allow for more customization of the graph. See [ST] **stcox diagnostics**.
11. Existing command `cc` has new option `tarone` that produces Tarone's (1985) adjustment of the Breslow–Day test for homogeneity of odds ratios. See [ST] **epitab**.
12. Existing command `stdes` has been renamed to `stdescribe`. `stdes` continues to work. See [ST] **stdescribe**.
13. The manual has an expanded glossary.

For a complete list of all the new features in Stata 10, see [U] **1.3 What's new**.

## References

- Freedman, L. S. 1982. Tables of the number of patients required in clinical trials using the logrank test. *Statistics in Medicine* 1: 121–129.
- Lachin, J. M., and M. A. Foulkes. 1986. Evaluation of sample size and power for analysis of survival with allowance for nonuniform patient entry, losses to follow-up, noncompliance, and stratification. *Biometrics* 42: 507–519.
- Rubinstein, L. V., M. H. Gail, and T. J. Santner. 1981. Planning the duration of a comparative clinical trial with loss to follow-up and a period of continued observation. *Journal of Chronic Diseases* 34: 469–479.
- Schoenfeld, D. 1981. The asymptotic properties of nonparametric tests for comparing survival distributions. *Biometrika* 68: 316–319.
- Tarone, R. E. 1985. On heterogeneity tests based on efficient scores. *Biometrika* 72: 91–95.

## Also See

[U] **1.3 What's new**

[R] **intro** — Introduction to base reference manual