Description

power cox computes sample size, power, or effect size for survival analyses that use Cox proportional hazards (PH) models. The results are obtained for the test of the effect of one covariate (binary or continuous) on time to failure adjusted for other predictors in a PH model. Effect size can be expressed as a regression coefficient (or log hazard-ratio) or as a hazard ratio. The command can account for the dependence between the covariate of interest and other model covariates, and it can adjust computations for censoring and for withdrawal of subjects for the study.

Quick start

Sample size for a test of $H_0 : \beta_1 = 0$ versus $H_a : \beta_1 \neq 0$ of a binary covariate in a Cox PH model given alternative coefficient *b1* of 0.4, no censoring, and default power of 0.8 and significance level $\alpha = 0.05$

    power cox .4

Same as above, specified using hazard ratio of 1.492

    power cox, hratio(1.492)

Sample size for a test of a continuous independent variable with *b1* of 1.2 and standard deviation of 0.47

    power cox 1.2, sd(.47)

Same as above, but for a model with multiple covariates and a squared multiple-correlation coefficient $R^2 = 0.2$ when the covariate of interest is regressed on all other covariates

    power cox 1.2, sd(.47) r2(.2)

Sample size and estimated number of events for a study with censoring, assuming an 80% event rate

    power cox .4, eventprob(.8)

Plot required sample size versus coefficient values of 0.2, 0.3, 0.4, 0.5, and 0.6

    power cox (.2(.1).6), graph

Sample size assuming 12% dropout of subjects in the study

    power cox .4, wdprob(.12)

Power of a test of a binary independent variable given sample size of 200

    power cox .4, n(200)

Power of a one-sided test of a continuous independent variable with *b1* of 1.2 and standard deviation of 0.47

    power cox 1.2, sd(.47) n(200) onesided

Effect size given power of 0.8 and sample sizes of 120, 130, 140, 150, and 160

    power cox, power(.8) n(120(10)160)

Menu

Statistics > Power, precision, and sample size

Syntax

Compute sample size

> power cox $\big[$ *b1* $\big]$ $\big[$, $\underline{\text{power}}$(*numlist*) *options* $\big]$

Compute power

> power cox $\big[$ *b1* $\big]$, n(*numlist*) $\big[$ *options* $\big]$

Compute effect size (target regression coefficient)

> power cox, n(*numlist*) $\underline{\text{power}}$(*numlist*) $\big[$ *options* $\big]$

where *b1* is the hypothesized regression coefficient (effect size) of a covariate of interest in a Cox PH model desired to be detected by a test with a prespecified power. *b1* may be specified either as one number or as a list of values in parentheses (see [U] **11.1.8 numlist**).

options	Description
Main	
* <u>a</u>lpha(*numlist*)	significance level; default is alpha(0.05)
* <u>pow</u>er(*numlist*)	power; default is power(0.8)
* <u>b</u>eta(*numlist*)	probability of type II error; default is beta(0.2)
* n(*numlist*)	sample size; required to compute power or effect size
<u>nfrac</u>tional	allow fractional sample size
* <u>hr</u>atio(*numlist*)	hazard ratio (exponentiated *b1*) associated with a one-unit increase in covariate of interest; specify instead of the regression coefficient *b1*; default is hratio(0.5)
* sd(*numlist*)	standard deviation of covariate of interest; default is sd(0.5)
* r2(*numlist*)	squared coefficient of multiple correlation with other covariates; default is r2(0)
* <u>eventp</u>rob(*numlist*)	overall probability of an event (failure) of interest; default is eventprob(1), meaning no censoring
* <u>failp</u>rob(*numlist*)	synonym for eventprob()
* <u>wdp</u>rob(*numlist*)	proportion of subjects anticipated to withdraw from the study; default is wdprob(0)
<u>eff</u>ect(*effect*)	specify the type of effect to display; default is effect(coefficient)
<u>dir</u>ection(<u>lower</u>|<u>upper</u>)	direction of the effect for effect-size determination; default is direction(lower), which means that the postulated value of the parameter is smaller than the hypothesized value
<u>one</u>sided	one-sided test; default is two sided
<u>par</u>allel	treat number lists in starred options or in command arguments as parallel when multiple values per option or argument are specified (do not enumerate all possible combinations of values)
Table	
[<u>no</u>]<u>table</u>[(*tablespec*)]	suppress table or display results as a table; see [PSS-2] **power, table**
<u>sav</u>ing(*filename* [, replace])	save the table data to *filename*; use replace to overwrite existing *filename*
Graph	
<u>graph</u>[(*graphopts*)]	graph results; see [PSS-2] **power, graph**
<u>notitle</u>	suppress the title

*Specifying a list of values in at least two starred options, or at least two command arguments, or at least one starred option and one argument results in computations for all possible combinations of the values; see [U] **11.1.8 numlist**. Also see the parallel option.

collect is allowed; see [U] **11.1.10 Prefix commands**.

notitle does not appear in the dialog box.

effect	Description
<u>coef</u>ficient	regression coefficient, *b1*; the default
<u>hr</u>atio	hazard ratio, exp(*b1*)
<u>lnhr</u>atio	log hazard-ratio; synonym for coefficient

where *tablespec* is

column[:label] [column[:label] [. . .]] [, tableopts]

column is one of the columns defined below, and *label* is a column label (may contain quotes and compound quotes).

column	Description	Symbol
alpha	significance level	α
power	power	$1 - \beta$
beta	type-II-error probability	β
N	number of subjects	N
delta	effect size	δ
E	total number of events (failures)	E
b1	regression coefficient	β_{1a}
hratio	hazard ratio	Δ
lnhratio	log hazard-ratio	$\ln(\Delta)$
sd	standard deviation	σ
R2	squared multiple-correlation coefficient	R^2
Pr_E	overall probability of an event (failure)	p_E
Pr_w	probability of withdrawals	p_w
target	target parameter; synonym for b1	
_all	display all supported columns	

Column beta is shown in the default table in place of column power if option beta() is specified.

Column b1 is shown in the default table in place of column hratio when a regression coefficient is specified.

Columns R2 and Pr_w are shown in the default table only if specified.

Options

Main

alpha(), power(), beta(), n(), nfractional; see [PSS-2] **power**. The nfractional option is allowed only for sample-size determination.

hratio(*numlist*) specifies the hazard ratio (or exponentiated regression coefficient) associated with a one-unit increase in the covariate of interest when other covariates are held constant. This value defines an effect size or the minimal clinically significant effect of a covariate on the response to be detected by a test with a certain power in a Cox PH model.

You can specify an effect size either as the regression coefficient *b1*, which is the command argument, or as the hazard ratio in hratio(). The default is hratio(0.5). If you specify hratio(#), the regression coefficient is computed as $b1 = \ln(\#)$. If you specify a regression coefficient *b1*, the hazard ratio is computed as $\exp(b1)$.

This option is not allowed with the effect-size determination.

sd(*numlist*) specifies the standard deviation of the covariate of interest. The default is sd(0.5).

r2(*numlist*) specifies the squared multiple-correlation coefficient between the covariate of interest and other predictors in a Cox PH model. The default is r2(0), meaning that the covariate of interest is independent of other covariates. This option defines the proportion of variance explained by the regression of the covariate of interest on other covariates used in the Cox model (see [R] **regress**).

eventprob(*numlist*) specifies the overall probability of a subject experiencing an event of interest (or failing, or not being censored) in the study. The default is eventprob(1), meaning that all subjects experience an event (or fail) in the study; that is, no censoring of subjects occurs.

failprob(*numlist*) is a synonym for eventprob().

wdprob(*numlist*) specifies the proportion of subjects anticipated to withdraw from a study. The default is wdprob(0). wdprob() is allowed only with sample-size computation.

effect(*effect*) specifies the type of the effect size to be reported in the output as delta. *effect* is one of coefficient, hratio, or lnhratio. By default, the effect size delta is the regression coefficient, effect(coefficient).

direction(), onesided, parallel; see [PSS-2] **power**. direction(lower) is the default.

___Table___

table, table(), notable; see [PSS-2] **power, table**.

saving(); see [PSS-2] **power**.

___Graph___

graph, graph(); see [PSS-2] **power, graph**. Also see the *column* table for a list of symbols used by the graphs.

The following option is available with power cox but is not shown in the dialog box:

notitle; see [PSS-2] **power**.

Remarks and examples

Remarks are presented under the following headings:

This entry describes the power cox command and the methodology for power and sample-size analysis for survival analyses that use Cox PH models. See [PSS-2] **Intro (power)** for a general introduction to power and sample-size analysis, and see [PSS-2] **power** for a general introduction to the power command using hypothesis tests. See *Survival data* in [PSS-2] **Intro (power)** for an introduction to power and sample-size analysis for survival data.

Introduction

Consider a survival study for which the goal is to investigate the effect of a covariate of interest, x_1, on time to failure, possibly adjusted for other predictors, $x_2, x_3, \ldots, x_p$, using the Cox PH model (Cox 1972). The effect is commonly measured as a hazard ratio Δ associated with a one-unit increase in x_1 when the other covariates $x_2, x_3, \ldots, x_p$ are held constant. For a binary predictor x_1, hazard ratio Δ corresponds to the two categories of x_1 when the other covariates are held constant.

In a Cox PH model, the hazard function is assumed to be

$$h(t) = h_0(t)\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p)$$

where no distributional assumption is made about the baseline hazard $h_0(t)$. Under this assumption, the regression coefficient β_1 is the log hazard-ratio $\ln(\Delta)$ associated with a one-unit increase in x_1 when the other predictors are held constant, and the exponentiated regression coefficient $\exp(\beta_1)$ is the hazard ratio Δ. Therefore, the effect of x_1 on time to failure can be investigated by performing an appropriate test based on the partial likelihood (Hosmer, Lemeshow, and May 2008; Klein and Moeschberger 2003) for the regression coefficient β_1 from a Cox model. Negative values of β_1 correspond to the reduction in hazard for a one-unit increase in x_1, and, conversely, positive values correspond to the increase in hazard for a one-unit increase in x_1.

power cox provides power and sample-size analysis for a test of the regression coefficient β_1 in a Cox model with the null hypothesis H_0: $(\beta_1, \beta_2, \ldots, \beta_p) = (0, \beta_2, \ldots, \beta_p)$ against the alternative H_a: $(\beta_1, \beta_2, \ldots, \beta_p) = (\beta_{1a}, \beta_2, \ldots, \beta_p)$. The postulated or null coefficient value is zero and a hypothesized or alternative coefficient value, specified as *b1* with power cox, is β_{1a}. (Similarly, Δ_a is a hypothesized or alternative value of the hazard ratio specified in the hratio() option.) The methods used are derived for the score test of H_0 versus H_a. In practice, however, the obtained results may be used in the context of the Wald test as well because the two tests usually lead to the same conclusions about the significance of the regression coefficient. Refer to *The conditional versus unconditional approaches* in [PSS-2] **power exponential** for more details about the results based on conditional and unconditional tests. From now on, we will refer to H_a as H_a: $\beta_1 = \beta_{1a}$ for simplicity.

power cox implements the method of Hsieh and Lavori (2000) for the sample-size and power computation, which reduces to the method of Schoenfeld (1983) for a binary covariate. The sample size is related to the power of a test through the number of events observed in the study; that is, for a fixed number of events, the power of a test is independent of the sample size. As a result, the sample size is estimated as the number of events divided by the overall probability of a subject failing in the study.

You can use power cox to

- compute required number of events and sample size when you know power and effect size expressed as a hazard ratio or regression coefficient (log hazard-ratio);

- compute power when you know sample size (number of events) and effect size expressed as a hazard ratio or regression coefficient (log hazard-ratio); or

- compute effect size and display it as a regression coefficient, a log hazard-ratio, or a hazard ratio when you know sample size (number of events) and power.

You can also account for the dependence between the covariate of interest and other model covariates in your analysis, adjust results for censoring, and adjust results for withdrawal of subjects from the study.

Using power cox

power cox computes sample size, power, or effect size for a test of one regression coefficient in a Cox PH model, holding coefficients of the other covariates constant. All computations are performed for a two-sided hypothesis test where, by default, the significance level is set to 0.05. You may change the significance level by specifying the alpha() option. You can specify the onesided option to request a one-sided test.

To compute sample size, you specify an effect size and, optionally, power of the test in the `power()` option. The default power is set to 0.8. By default, the computed sample size is rounded up. You can specify the `nfractional` option to see the corresponding fractional sample size; see *Fractional sample sizes* in [PSS-4] **Unbalanced designs** for an example. The `nfractional` option is allowed only for sample-size determination.

To compute power, you must specify the sample size in the `n()` option and an effect size.

An effect size may be specified either as a regression coefficient supplied as the command argument *b1* or as a hazard ratio supplied in the `hratio()` option. If neither is specified, a hazard ratio of 0.5 is assumed.

To compute effect size, which may be expressed either as a regression coefficient (log hazard-ratio) or hazard ratio, you must specify the sample size in the `n()` option; the power in the `power()` option; and, optionally, the direction of the effect. The direction is lower by default, `direction(lower)`, which means, for example, that the target regression coefficient is assumed to be negative. This is equivalent to the hazard ratio being less than one. You can change the direction to upper, which means that the target regression coefficient is assumed to be positive, by specifying the `direction(upper)` option. This is equivalent to the hazard ratio being greater than one.

As we mentioned above, the effect size for `power cox` may be expressed as a regression coefficient or, equivalently, a log hazard-ratio, or as a hazard ratio. By default, the effect size, which is labeled as `delta` in the output, corresponds to the regression coefficient. You can change this by specifying the `effect()` option: `effect(coefficient)` (the default) reports the regression coefficient, `effect(hratio)` reports the hazard ratio, and `effect(lnhratio)` reports the log hazard-ratio.

The standard deviation of the covariate of interest is set to 0.5 by default and may be changed by specifying the `sd()` option. In the presence of additional covariates in a Cox model, you can use the `r2()` option to specify the correlation between the covariate of interest and other covariates in the model.

All computations assume no censoring. In the presence of censoring, you can use the `eventprob()` option to specify an overall probability of an event or failure. When computing sample size, you can also adjust for withdrawal of subjects from the study by specifying the anticipated proportion of withdrawals in the `wdprob()` option.

In the following sections, we describe the use of `power cox` accompanied by examples for computing sample size, power, and effect size.

Computing sample size

To compute sample size and number of events, you must specify an effect size (a regression coefficient or a hazard ratio) and, optionally, the power of the test in the `power()` option. A default power of 0.8 is assumed if `power()` is not specified. A hazard ratio of 0.5 is assumed if an effect size is not specified.

Computing sample size in the absence of censoring

First, consider a type I study in which all subjects fail by the end of the study (no censoring). Then the required sample size is the same as the number of events required to be observed in a study.

▷ Example 1: Sample size for a model with a binary covariate of interest

Consider a survival study for which the goal is to investigate the effect of a treatment on survival times of subjects. The covariate of interest is binary with levels defining whether a subject receives the treatment (the experimental group) or a placebo (the control or placebo group). Prior to conducting the study, investigators need an estimate of the sample size that ensures that a ratio of hazards of the experimental group to the control group of 0.5 ($\beta_{1a} = \ln(0.5) = -0.6931$) can be detected with a power of 80% with a two-sided, 5%-level test. Under 1:1 randomization, a subject has a 50% chance of receiving the treatment. The corresponding binary covariate follows a Bernoulli distribution with the probability of a subject receiving a treatment, p, equal to 0.5. As such, the standard deviation of the covariate is $\{p(1-p)\}^{1/2} = 0.5$. Because these study parameters correspond to default values of power cox, to obtain the sample size for the above study, we simply type

```
. power cox
Estimated sample size for Cox PH regression
Wald test
H0: beta1 = 0  versus  Ha: beta1 != 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =   -0.6931  (coefficient)
       hratio =    0.5000
           sd =    0.5000
Censoring:
         Pr_E =    1.0000
Estimated number of events and sample size:
            E =        66
            N =        66
```

Recall that, by default, a hazard ratio of 0.5, corresponding to the regression coefficient of $\ln(0.5) = -0.6931$, is assumed. From the output, we see that 66 events (failures) are required to be observed in the study to ensure a power of 80% to detect an alternative H_a: $\beta_1 = -0.6931$ using a two-sided test with a 0.05 significance level. Because we have no censoring (Pr_E = 1.0000), a total of 66 subjects is needed in the study to observe 66 events.

◁

▷ Example 2: Alternative ways of specifying effect

In example 1, the effect size delta corresponds to the regression coefficient, the default. We can use the effect(hratio) option to redefine delta as a hazard ratio.

```
. power cox, effect(hratio)

Estimated sample size for Cox PH regression
Wald test
H0: beta1 = 0  versus  Ha: beta1 != 0

Study parameters:

        alpha =    0.0500
        power =    0.8000
        delta =    0.5000  (hazard ratio)
       hratio =    0.5000
           sd =    0.5000

Censoring:

         Pr_E =    1.0000

Estimated number of events and sample size:

            E =        66
            N =        66
```

We can also obtain the same results as in example 1 by directly specifying the value of the coefficient.

```
. power cox -0.6931

Estimated sample size for Cox PH regression
Wald test
H0: beta1 = 0  versus  Ha: beta1 != 0

Study parameters:

        alpha =    0.0500
        power =    0.8000
        delta =   -0.6931  (coefficient)
           b1 =   -0.6931
           sd =    0.5000

Censoring:

         Pr_E =    1.0000

Estimated number of events and sample size:

            E =        66
            N =        66
```

The specified regression coefficient, b1, is reported in the output instead of the hazard ratio hratio.

◁

Suppose now that the covariate of interest, x_1, is continuous. Hsieh and Lavori (2000) extend the formula of Schoenfeld (1983) for the number of events to the case when a covariate is continuous. They also relax the assumption of Schoenfeld (1983) about the independence of x_1 of other covariates and provide an adjustment to the estimate of the number of events for possible correlation.

▷ Example 3: Sample size for a model with a continuous covariate of interest

Consider an example from Hsieh and Lavori (2000) of a study of multiple-myeloma patients treated with alkylating agents (Krall, Uthoff, and Harley 1975). Although in the original study of multiple-myeloma patients, 17 of a total of 65 patients are censored; here we assume that all patients die by the end of the study (a type I study, no censoring). Suppose that the covariate of interest, x_1, is the log of the amount of blood urea nitrogen (BUN) measured in a patient. The sample size for a one-sided, 5%-level test to detect a coefficient (log hazard-ratio) of 1 for a unit increase in x_1 with a power of 80% is required. The standard deviation of x_1 is 0.3126. To obtain an estimate of the sample size, we supply *b1*, 1, as an argument, the sd(0.3126) option, and the onesided option to request a one-sided test.

```
. power cox 1, sd(0.3126) onesided
Estimated sample size for Cox PH regression
Wald test
H0: beta1 = 0  versus  Ha: beta1 > 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =    1.0000  (coefficient)
           b1 =    1.0000
           sd =    0.3126
Censoring:
         Pr_E =    1.0000
Estimated number of events and sample size:
            E =        64
            N =        64
```

The estimate of the required number of events and the sample size is 64.

◁

Based on the derivation in Schoenfeld (1983) and Hsieh and Lavori (2000), sample-size estimates in the above examples may be used if other covariates are also present in the model as long as these covariates are independent of the covariate of interest. The independence assumption holds for randomized studies, but it is not true for nonrandomized studies often encountered in practice. Also, in many studies, the main covariate of interest will often be correlated with other covariates. For example, age and gender will often be confounded with the covariate of interest, such as smoking. Below we investigate the effect of the confounding factor on the estimate of the required number of events.

▷ Example 4: Sample size when covariates are not independent

Continuing with example 3, suppose that we want to adjust the effect of the covariate BUN for eight other covariates in the model. Hsieh and Lavori (2000) report the coefficient of determination of $R^2 = 0.1837$ from regression of the log of BUN, x_1, on the eight other covariates. We specify this value in the r2() option.

```
. power cox 1, sd(0.3126) onesided r2(0.1837)
Estimated sample size for Cox PH regression
Wald test
H0: beta1 = 0  versus  Ha: beta1 > 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =    1.0000  (coefficient)
           b1 =    1.0000
           sd =    0.3126
           R2 =    0.1837
Censoring:
         Pr_E =    1.0000
Estimated number of events and sample size:
            E =        78
            N =        78
```

The number of events required to be observed in the study increases from 64 to 78 because of the correlation between BUN and the other covariates. The new estimate is equal to the original estimate multiplied by the variance inflation factor $\text{VIF} = 1/(1-R^2)$. Likewise, the number of subjects increases from 64 to 78.

◁

Computing sample size in the presence of censoring

In the previous section, we assumed that all subjects experience an event by the end of the study. In practice, the study often terminates after a fixed time, T. As a result, some subjects may not experience an event by the end of the study (a type II study). These subjects are censored. To obtain an estimate of the sample size in the presence of censoring, an estimate of the overall probability of a subject not being censored is required. The investigator may already have such an estimate from previous studies, or this probability may be computed as suggested in the literature (Schoenfeld [1983], Lachin and Foulkes [1986], Barthel et al. [2006], and Barthel, Royston, and Babiker [2005], also see [PSS-2] **power logrank** and [PSS-2] **power exponential**).

▷ Example 5: Sample size in the presence of censoring

Consider the study from example 3. In reality, as mentioned earlier, 17 of a total of 65 patients survived until the end of the study. The overall deathrate is estimated as $1 - 17/65 = 0.738$. We specify the overall probability of an event in the eventprob() option.

```
. power cox 1, sd(0.3126) onesided eventprob(0.738)
Estimated sample size for Cox PH regression
Wald test
H0: beta1 = 0  versus  Ha: beta1 > 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =    1.0000  (coefficient)
           b1 =    1.0000
           sd =    0.3126
Censoring:
         Pr_E =    0.7380
Estimated number of events and sample size:
            E =        64
            N =        86
```

In the presence of censoring, the number of subjects required in the study increases from 64 to 86. The number of events remains the same (64) because the only change in the study is the presence of censoring, and censoring is assumed to be independent of failure (event) times.

◁

▷ Example 6: Sample size in the presence of censoring adjusting for other covariates

Continuing with example 5, if we also adjust for the correlation between the log of BUN and other covariates, we obtain the estimate of the sample size to be 106.

```
. power cox, hratio(2.7182) sd(0.3126) onesided eventprob(0.738) r2(0.1837)
Estimated sample size for Cox PH regression
Wald test
H0: beta1 = 0  versus  Ha: beta1 > 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =    1.0000  (coefficient)
       hratio =    2.7182
           sd =    0.3126
           R2 =    0.1837
Censoring:
         Pr_E =    0.7380
Estimated number of events and sample size:
            E =        78
            N =       106
```

In the above example, for no other reason than variety, rather than supplying the coefficient of 1, we used the hratio() option to specify the size of the effect expressed as the hazard ratio $\exp(1) = 2.7182$.

◁

❑ Technical note

Supplying the coefficient (log hazard-ratio) of 1 or -1 [or, respectively, the hazard ratio of $\exp(1) = 2.7182$ or $\exp(-1) = 1/2.7182 = 0.36788$] is irrelevant for sample-size and power determination because it results in the same estimates of sample size and power. However, the sign of the coefficient (or the value of the hazard ratio being larger or smaller than one) is important at the analysis stage because it determines the direction of the effect associated with a one-unit increase of a covariate value.

❑

Often, in practice, subjects may withdraw from a study before it terminates. As a result, the information about the subjects' response is lost. The proportion of subjects anticipated to withdraw from a study may be specified by using wdprob(). Refer to *Survival data* in [PSS-2] **Intro (power)** and *Withdrawal of subjects from the study* in [PSS-2] **power logrank** for a more detailed description and an example.

Link to the sample-size and power computation for the log-rank test

The score test of the regression coefficient of a binary covariate in a Cox model with one binary predictor is the same (in the absence of tied observations) as the log-rank test comparing survivor functions of two groups defined by this covariate. Powers of the two tests are the same and so are the formulas for the number of events (Schoenfeld 1983; Schoenfeld 1981). As such, the required number of events computed by power cox for the Cox model and the Schoenfeld method of power logrank for the log-rank test are the same.

▷ Example 7: Using power logrank for a binary covariate

Using power logrank (see [PSS-2] **power logrank**) for the study described in example 1 yields the same estimates of 66 for both the required number of events and the required sample size.

```
. power logrank, schoenfeld
Estimated sample sizes for two-sample comparison of survivor functions
Log-rank test, Schoenfeld method
H0: ln(HR) = 0  versus  Ha: ln(HR) != 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =   -0.6931  (log hazard-ratio)
       hratio =    0.5000
Censoring:
         Pr_E =    1.0000
Estimated number of events and sample sizes:
            E =        66
            N =        66
  N per group =        33
```

◁

Schoenfeld (1983) demonstrates that the same formula can also be used to compute the required number of events when other covariates are present in a Cox model, provided that the binary covariate of interest is independent of these covariates and that the covariates are not extremely unbalanced. Although the formulas for the number of events are the same whether covariates are present or not, it is important to adjust for covariates when analyzing the data to avoid loss of power; see Schoenfeld (1983) for details.

Væth and Skovlund (2004) demonstrate that for a continuous covariate of interest in a Cox model, the sample-size formula for the log-rank test with the value of the hazard ratio equal to $\exp(2\beta_{1a}\sigma)$ and with equal-group allocation may be used to obtain the required sample size. Indeed, by substituting the above expression for the hazard ratio into the sample-size formula for the log-rank test, one obtains the sample-size formula derived in Hsieh and Lavori (2000) for a Cox model.

▷ Example 8: Using power logrank for a continuous covariate

For example, we obtain the same estimate of the total number of events as computed in example 3 by using power logrank with the schoenfeld option and with the value of the hazard ratio equal to $\exp(2\beta_{1a}\sigma) = \exp(2 \times 1 \times 0.3126) = 1.8686$.

```
. power logrank, hratio(1.8686) onesided schoenfeld
Estimated sample sizes for two-sample comparison of survivor functions
Log-rank test, Schoenfeld method
H0: ln(HR) = 0  versus  Ha: ln(HR) > 0
Study parameters:
        alpha =    0.0500
        power =    0.8000
        delta =    0.6252  (log hazard-ratio)
       hratio =    1.8686
Censoring:
         Pr_E =    1.0000
Estimated number of events and sample sizes:
            E =        64
            N =        64
  N per group =        32
```

◁

Computing power

To compute power, you must specify the sample size in the `n()` option and an effect size (a regression coefficient or a hazard ratio). A hazard ratio of 0.5 is assumed if an effect size is not specified.

▷ Example 9: Power determination

In example 6, the required number of patients was estimated to be 106 to ensure a power of 80% for a 0.05 one-sided test to detect a value of 1 in the regression coefficient. Suppose that we can recruit only 65 subjects. How does this reduction in sample size affect the power of the test to detect the alternative $H_a\colon \beta_1 = 1$?

```
. power cox 1, sd(0.3126) onesided r2(0.1837) eventprob(0.738) n(65)
Estimated power for Cox PH regression
Wald test
H0: beta1 = 0   versus   Ha: beta1 > 0
Study parameters:
        alpha =     0.0500
            N =         65
        delta =     1.0000  (coefficient)
           b1 =     1.0000
           sd =     0.3126
           R2 =     0.1837
Number of events and censoring:
            E =         48
         Pr_E =     0.7380
Estimated power:
        power =     0.6222
```

When the sample size decreases from 106 to 65, the corresponding number of events decreases from 78 to 48, and power decreases from 80% to 62%.

◁

▷ Example 10: Multiple values of study parameters

Suppose we want to investigate the effect of a correlation between a covariate of interest and other model covariates on power. Continuing with example 9, we can specify a list (see [U] 11.1.8 numlist) of correlations in the `r2()` option.

```
. power cox 1, sd(0.3126) onesided r2(0.1(0.1)0.5) eventprob(0.738) n(65)
Estimated power for Cox PH regression
Wald test
H0: beta1 = 0   versus   Ha: beta1 > 0
```

alpha	power	N	E	delta	b1	sd	R2	Pr_E
.05	.6588	65	48	1	1	.3126	.1	.738
.05	.6147	65	48	1	1	.3126	.2	.738
.05	.5662	65	48	1	1	.3126	.3	.738
.05	.5128	65	48	1	1	.3126	.4	.738
.05	.4547	65	48	1	1	.3126	.5	.738

As the correlation increases, the power decreases.

For multiple values of parameters, the results are automatically displayed in a table, as we see above. For more examples of tables, see [PSS-2] **power, table**. If you wish to produce a power plot, see [PSS-2] **power, graph**.

◁

Computing effect size

Effect size δ for a test of a coefficient in a Cox PH model is defined as a coefficient (or, equivalently, a log hazard-ratio) or a hazard ratio, corresponding to a one-unit change in the tested covariate holding other model covariates constant.

Sometimes, we may be interested in determining the smallest effect that yields a statistically significant result for prespecified sample size and power. In this case, both power and sample size must be specified in options power() and n(), respectively. Additionally, you may also choose the direction of the effect by specifying the direction() option. direction(lower), the default, assumes $\beta_{1a} < 0$ (or $\Delta_a < 1$), corresponding to the reduction in hazard for a unit change in a covariate. direction(upper) assumes $\beta_{1a} > 0$ (or $\Delta_a > 1$).

▷ Example 11: Effect-size determination

Continuing with example 9, if a power of 62% is unacceptable to investigators, they may want to determine the smallest value of the regression coefficient that can be detected with a preserved power of 80%. To obtain this estimate, we specify both the n() and the power() options.

```
. power cox, sd(0.3126) onesided r2(0.1837) eventprob(0.738) n(65) power(0.8)
> direction(upper)

Estimated target coefficient for Cox PH regression
Wald test
H0: beta1 = 0   versus   Ha: beta1 > 0

Study parameters:

        alpha =    0.0500
        power =    0.8000
            N =        65
           sd =    0.3126
           R2 =    0.1837

Number of events and censoring:

            E =        48
         Pr_E =    0.7380

Estimated effect size and target coefficient:

        delta =    1.2711   (coefficient)
           b1 =    1.2711
```

With only 65 subjects, the smallest change in a coefficient (log hazards) for a one-unit increase in the log of BUN, which can be detected with a preserved 80% power, is roughly 1.27, corresponding to a 27% increase in the coefficient of 1 desired to be detected originally in example 6.

◁

Performing analyses using a Cox PH model

After the data are collected, one can use stcox and test to fit the Cox PH model and perform a Wald test, as we demonstrate below.

▷ Example 12: Performing a Wald test

We demonstrate how to perform a Wald test for the regression coefficient of the log of BUN from a Cox model using the data from Krall, Uthoff, and Harley (1975) described in example 3. myeloma.dta consists of 11 variables, described below.

```
. use https://www.stata-press.com/data/r19/myeloma
(Multiple myeloma patients)

. describe

Contains data from https://www.stata-press.com/data/r19/myeloma.dta
 Observations:            65                          Multiple myeloma patients
    Variables:            11                          11 Feb 2024 19:26
```

Variable name	Storage type	Display format	Value label	Variable label
time	float	%9.0g		Survival time from diagnosis to nearest month + 1
died	byte	%9.0g		0 – Alive, 1 – Dead
lnbun	float	%9.0g		Log BUN at diagnosis
hemo	float	%9.0g		Hemoglobin at diagnosis
platelet	byte	%9.0g	normal	Platelets at diagnosis
age	byte	%9.0g		Age (complete years)
lnwbc	float	%9.0g		Log WBC at diagnosis
fracture	byte	%9.0g	present	Fractures at diagnosis
lnbm	float	%9.0g		Log % of plasma cells in bone marrow
protein	byte	%9.0g		Proteinuria at diagnosis
scalcium	byte	%9.0g		Serum calcium (mgm%)

```
Sorted by:
```

Before using stcox to fit a Cox model, we need to set up the data by using stset (see [ST] **stset**). The analysis-time variable is time, and the failure variable is died.

```
. stset time, failure(died)

Survival-time data settings

         Failure event: died!=0 & died<.
Observed time interval: (0, time]
    Exit on or before: failure
──────────────────────────────────────────────────────────────────
        65  total observations
         0  exclusions
──────────────────────────────────────────────────────────────────
        65  observations remaining, representing
        48  failures in single-record/single-failure data
   1,560.5  total analysis time at risk and under observation
                                  At risk from t =          0
                            Earliest observed entry t =      0
                                Last observed exit t =      92
```

We include all nine covariates in the model and perform a fit by using stcox. Then we perform a Wald test of $H_0: \beta_1 = 1$ for the coefficient of lnbun using test.

```
. stcox lnbun hemo platelet age lnwbc fracture lnbm protein scalcium, nohr

        Failure _d: died
  Analysis time _t: time

Iteration 0:  Log likelihood = -154.85799
Iteration 1:  Log likelihood = -146.68114
Iteration 2:  Log likelihood = -146.29446
Iteration 3:  Log likelihood = -146.29404
Refining estimates:
Iteration 0:  Log likelihood = -146.29404

Cox regression with Breslow method for ties

No. of subjects =      65                        Number of obs =       65
No. of failures =      48
Time at risk    = 1,560.5
                                                 LR chi2(9)    =   17.13
Log likelihood = -146.29404                      Prob > chi2   =  0.0468
```

_t	Coefficient	Std. err.	z	P>|z|	[95% conf. interval]	
lnbun	1.798354	.6483293	2.77	0.006	.5276519	3.069056
hemo	-.1263119	.0718333	-1.76	0.079	-.2671026	.0144789
platelet	-.2505915	.5074656	-0.49	0.621	-1.245206	.7440228
age	-.0127949	.019475	-0.66	0.511	-.0509653	.0253755
lnwbc	.3537259	.7131935	0.50	0.620	-1.044108	1.75156
fracture	.3378767	.4072774	0.83	0.407	-.4603722	1.136126
lnbm	.3589346	.4860298	0.74	0.460	-.5936663	1.311535
protein	.0130672	.0261696	0.50	0.618	-.0382243	.0643587
scalcium	.1259479	.1034015	1.22	0.223	-.0767153	.3286112

```
. test lnbun = 1

 ( 1)  lnbun = 1

           chi2(  1) =    1.52
         Prob > chi2 =  0.2182
```

By default, `stcox` reports estimates of hazard ratios and the two-sided tests of the equality of a coefficient to zero. We use the `nohr` option to request estimates of coefficients. From the output table, a one-sided test of $H_0: \beta_1 = 0$ versus $H_a: \beta_1 > 0$ is rejected at a 0.05 level (one-sided p-value is $0.006/2 = 0.003 < 0.05$). The estimate of the log-hazard difference associated with a one-unit increase of `lnbun` is $\hat{\beta}_1 = 1.8$. From the `test` output, we cannot reject the hypothesis of $H_0: \beta_1 = 1$.

For these data, the observed effect size (coefficient) of 1.8 is large enough for the sample size of 65 to be sufficient to reject the null hypothesis of no effect of the BUN on the survival of subjects ($H_0: \beta_1 = 0$). However, if the goal of the study were to ensure that the test detects the effect size corresponding to the coefficient of at least 1 with 80% power, a sample of approximately 106 subjects would have been required.

◁

Stored results

power cox stores the following in r():

Scalars

r(alpha)	significance level
r(power)	power
r(beta)	probability of a type II error
r(delta)	effect size
r(N)	sample size
r(nfractional)	1 if nfractional is specified, 0 otherwise
r(onesided)	1 for a one-sided test, 0 otherwise
r(E)	total number of events (failures)
r(hratio)	hazard ratio under the alternative hypothesis
r(b1)	regression coefficient under the alternative hypothesis
r(sd)	standard deviation
r(R2)	squared multiple correlation (if specified)
r(Pr_E)	probability of an event (failure) (if specified)
r(Pr_w)	proportion of withdrawals (if specified)
r(separator)	number of lines between separator lines in the table
r(divider)	1 if divider is requested in the table, 0 otherwise

Macros

r(type)	test
r(method)	cox
r(effect)	coefficient, hratio, or lnhratio
r(direction)	lower or upper
r(columns)	displayed table columns
r(labels)	table column labels
r(widths)	table column widths
r(formats)	table column formats

Matrices

r(pss_table)	table of results

Methods and formulas

Let β_1 denote the regression coefficient of the covariate of interest, x_1, from a Cox PH model, possibly in the presence of other covariates, $x_2, \ldots, x_p$; and let Δ denote the hazard ratio associated with a one-unit increase of x_1 when other covariates are held constant. Under the PH model, $\beta_1 = \ln(\Delta)$, where $\ln(\Delta)$ is the change in log hazards associated with a one-unit increase in x_1 when other covariates are held constant.

Define E and n to be the total number of events (failures) and the total number of subjects required in the study; σ to be the standard deviation of x_1; p_E to be the overall probability of an event (failure); R^2 to be the proportion of variance explained by the regression of x_1 on $x_2, \ldots, x_p$ (or squared multiple-correlation coefficient); p_w to be the proportion of subjects withdrawn from a study (lost to follow-up); α to be the significance level; β to be the probability of a type II error; and $z_{(1-\alpha/k)}$ and $z_{(1-\beta)}$ to be the $(1-\alpha/k)$th and the $(1-\beta)$th quantiles of the standard normal distribution, respectively, with $k = 1$ for the one-sided test and $k = 2$ for the two-sided test.

The total number of events required to be observed in a study to ensure a power of $1 - \beta$ of a test to detect the regression coefficient, β_1, with a significance level α, according to Hsieh and Lavori (2000), is

$$E = \frac{(z_{1-\alpha/k} + z_{1-\beta})^2}{\sigma^2 \beta_1^2 (1 - R^2)}$$

For the case of randomized study and a binary covariate x_1, this formula was derived in Schoenfeld (1983). The formula is an approximation and relies on a set of assumptions such as distinct failure times, all subjects completing the course of the study (no withdrawal), and a local alternative under which $\ln(\Delta)$ is assumed to be of order $O(n^{-1/2})$. The formula is derived for the score test but may be applied to other tests (Wald, for example) that are based on the partial likelihood of a Cox model because all of these tests are asymptotically equivalent (Schoenfeld 1983; Hosmer, Lemeshow, and May 2008; Klein and Moeschberger 2003).

The total sample size required to observe the total number of events, E, is given by

$$n = \frac{E}{p_E}$$

If the `nfractional` option is not specified, the computed sample size is rounded up.

To account for a proportion of subjects, p_w, withdrawn from a study, a conservative adjustment to the total sample size suggested in the literature (Freedman 1982; Machin and Campbell 2005) is applied as follows:

$$n_w = \frac{n}{1 - p_w}$$

Withdrawal is assumed to be independent of administrative censoring and failure (event) times.

Power is estimated using the formula

$$1 - \beta = \Phi\left[|\beta_1|\sigma\{np_E(1 - R^2)\}^{1/2} - z_{1-\alpha/k}\right]$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

The estimate of the regression coefficient for a fixed power, $1 - \beta$, and a sample size, n, is computed as

$$\beta_1^2 = \frac{(z_{1-\alpha/k} + z_{1-\beta})^2}{\sigma^2 np_E(1 - R^2)}$$

Either of the two values $|\beta_1|$ and $-|\beta_1|$ satisfy the above equation. By default or if `direction(lower)` is specified, `power cox` reports the negative of the two values, which corresponds to the reduction in a hazard of a failure for a one-unit increase in x_1. If `direction(upper)` is specified, `power cox` reports the positive of the two values, which corresponds to the increase in a hazard of a failure for a one-unit increase in x_1.

Similarly, if the `effect(hratio)` option is used, the corresponding value of the hazard ratio less than 1 for `direction(lower)` and greater than 1 for `direction(upper)` is reported to reflect, respectively, the reduction or increase in hazard for a one-unit increase in x_1.

References

Barthel, F. M.-S., A. G. Babiker, P. Royston, and M. K. B. Parmar. 2006. Evaluation of sample size and power for multi-arm survival trials allowing for non-uniform accrual, non-proportional hazards, loss to follow-up and cross-over. *Statistics in Medicine* 25: 2521–2542. https://doi.org/10.1002/sim.2517.

Barthel, F. M.-S., P. Royston, and A. G. Babiker. 2005. A menu-driven facility for complex sample size calculation in randomized controlled trials with a survival or a binary outcome: Update. *Stata Journal* 5: 123–129.

Cox, D. R. 1972. Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society*, B ser., 34: 187–220. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x.

Freedman, L. S. 1982. Tables of the number of patients required in clinical trials using the logrank test. *Statistics in Medicine* 1: 121–129. https://doi.org/10.1002/sim.4780010204.

Hosmer, D. W., Jr., S. A. Lemeshow, and S. May. 2008. *Applied Survival Analysis: Regression Modeling of Time to Event Data.* 2nd ed. New York: Wiley.

Hsieh, F. Y., and P. W. Lavori. 2000. Sample-size calculations for the Cox proportional hazards regression model with nonbinary covariates. *Controlled Clinical Trials* 21: 552–560. https://doi.org/10.1016/S0197-2456(00)00104-5.

Klein, J. P., and M. L. Moeschberger. 2003. *Survival Analysis: Techniques for Censored and Truncated Data.* 2nd ed. New York: Springer.

Krall, J. M., V. A. Uthoff, and J. B. Harley. 1975. A step-up procedure for selecting variables associated with survival. *Biometrics* 31: 49–57. https://doi.org/10.2307/2529709.

Lachin, J. M., and M. A. Foulkes. 1986. Evaluation of sample size and power for analyses of survival with allowance for nonuniform patient entry, losses to follow-up, noncompliance, and stratification. *Biometrics* 42: 507–519. https://doi.org/10.2307/2531201.

Machin, D., and M. J. Campbell. 2005. *Design of Studies for Medical Research.* Chichester, UK: Wiley.

Schoenfeld, D. A. 1981. The asymptotic properties of nonparametric tests for comparing survival distributions. *Biometrika* 68: 316–319. https://doi.org/10.2307/2335833.

———. 1983. Sample-size formula for the proportional-hazards regression model. *Biometrics* 39: 499–503. https://doi.org/10.2307/2531021.

Væth, M., and E. Skovlund. 2004. A simple approach to power and sample size calculations in logistic regression and Cox regression models. *Statistics in Medicine* 23: 1781–1792. https://doi.org/10.1002/sim.1753.

Also see [PSS-2] **Intro (power)** for more references.

Also see

[PSS-2] **power** — Power and sample-size analysis for hypothesis tests

[PSS-2] **power exponential** — Power analysis for a two-sample exponential test

[PSS-2] **power logrank** — Power analysis for the log-rank test

[PSS-2] **power, graph** — Graph results from the power command

[PSS-2] **power, table** — Produce table of results from the power command

[PSS-5] **Glossary**

[R] **test** — Test linear hypotheses after estimation

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

[ST] **sts test** — Test equality of survivor functions