Title

stata.com

lincom — Linear combinations of parameters

Description

`lincom` computes point estimates, standard errors, t or z statistics, p-values, and confidence intervals for linear combinations of coefficients after any estimation command, including survey estimation. Results can optionally be displayed as odds ratios, hazard ratios, incidence-rate ratios, or relative-risk ratios.

Quick start

Point estimate and confidence interval for sum of coefficients of `x1` and `x2`

```
lincom x1 + x2
```

Same as above, but report results as a relative-risk ratio

```
lincom x1 + x2, rrr
```

Same as above, but use coefficients from second equation of a multiequation model

```
lincom [2]x1 + [2]x2, rrr
```

Difference between coefficients of first and third level of categorical variable `a`

```
lincom 1.a - 3.a
```

Sum of coefficients of `x1` and `x2` after a model adjusted for complex survey design

```
lincom x1 + x2
```

Menu

Statistics > Postestimation

Syntax

```
lincom exp [ , options ]
```

options	Description
eform	generic label; exp(b)
or	odds ratio
hr	hazard ratio
shr	subhazard ratio
irr	incidence-rate ratio
rrr	relative-risk ratio
level(#)	set confidence level; default is level(95)
display_options	control column formats
df(#)	use t distribution with # degrees of freedom for computing p-values and confidence intervals

exp is any linear combination of coefficients that is a valid syntax for test; see [R] **test**. *exp* must not contain an equal sign.

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

df(#) does not appear in the dialog box.

Options

eform, or, hr, shr, irr, and rrr all report coefficient estimates as $\exp(\widehat{\beta})$ rather than $\widehat{\beta}$. Standard errors and confidence intervals are similarly transformed. or is the default after logistic. The only difference in these options is how the output is labeled.

Option	Label	Explanation	Example commands
eform	exp(b)	Generic label	cloglog
or	Odds ratio	Odds ratio	logistic, logit
hr	Haz. ratio	Hazard ratio	stcox, streg
shr	SHR	Subhazard ratio	stcrreg
irr	IRR	Incidence-rate ratio	poisson
rrr	RRR	Relative-risk ratio	mlogit

exp may not contain any additive constants when you use the eform, or, hr, irr, or rrr option.

level(#) specifies the confidence level, as a percentage, for confidence intervals. The default is level(95) or as set by set level; see [U] **20.8 Specifying the width of confidence intervals**.

display_options: cformat(%*fmt*), pformat(%*fmt*), and sformat(%*fmt*); see [R] **Estimation options**.

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

df(#) specifies that the t distribution with # degrees of freedom be used for computing p-values and confidence intervals. The default is to use e(df_r) degrees of freedom or the standard normal distribution if e(df_r) is missing.

Remarks and examples

stata.com

Remarks are presented under the following headings:

Using lincom

After fitting a model and obtaining estimates for coefficients $\beta_1, \beta_2, \ldots, \beta_k$, you may want to view estimates for linear combinations of the β_i, such as $\beta_1 - \beta_2$. `lincom` can display estimates for any linear combination of the form $c_0 + c_1\beta_1 + c_2\beta_2 + \cdots + c_k\beta_k$.

`lincom` works after any estimation command for which `test` works. Any valid expression for `test` syntax 1 (see [R] **test**) is a valid expression for `lincom`.

`lincom` is useful for viewing odds ratios, hazard ratios, etc., for one group (that is, one set of covariates) relative to another group (that is, another set of covariates). See the examples below.

▷ Example 1

We perform a linear regression:

```
. use https://www.stata-press.com/data/r18/regress
. regress y x1 x2 x3
```

Source	SS	df	MS	Number of obs	=	148
				F(3, 144)	=	96.12
Model	3259.3561	3	1086.45203	Prob > F	=	0.0000
Residual	1627.56282	144	11.3025196	R-squared	=	0.6670
				Adj R-squared	=	0.6600
Total	4886.91892	147	33.2443464	Root MSE	=	3.3619

y	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
x1	1.457113	1.07461	1.36	0.177	-.666934	3.581161
x2	2.221682	.8610358	2.58	0.011	.5197797	3.923583
x3	-.006139	.0005543	-11.08	0.000	-.0072345	-.0050435
_cons	36.10135	4.382693	8.24	0.000	27.43863	44.76407

To see the difference of the coefficients of `x2` and `x1`, we type

```
. lincom x2 - x1
 ( 1)  - x1 + x2 = 0
```

y	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
(1)	.7645682	.9950282	0.77	0.444	-1.20218	2.731316

The expression can be any linear combination.

```
. lincom 3*x1 + 500*x3
 ( 1)  3*x1 + 500*x3 = 0
```

y	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
(1)	1.301825	3.396624	0.38	0.702	-5.411858	8.015507

Nonlinear expressions are not allowed.

```
. lincom x2/x1
not possible with test
r(131);
```

For information about estimating nonlinear expressions, see [R] **nlcom**.

◁

❑ Technical note

`lincom` uses the same shorthands for coefficients as does `test` (see [R] **test**). When you type `x1`, for instance, `lincom` knows that you mean the coefficient of `x1`. The formal syntax for referencing this coefficient is actually `_b[x1]`, or alternatively, `_coef[x1]`. So, more formally, in the last example we could have typed

```
. lincom 3*_b[x1] + 500*_b[x3]
```

(*output omitted*)

❑

Odds ratios and incidence-rate ratios

After logistic regression, the `or` option can be specified with `lincom` to display odds ratios for any effect. Incidence-rate ratios after commands such as `poisson` can be similarly obtained by specifying the `irr` option.

▷ Example 2

Consider the low birthweight dataset from Hosmer, Lemeshow, and Sturdivant (2013, 24). We fit a logistic regression model of low birthweight (variable `low`) on the following variables:

Variable	Description	Coding
`age`	age in years	
`race`	race	1 if white, 2 if black, 3 if other
`smoke`	smoking status	1 if smoker, 0 if nonsmoker
`ht`	history of hypertension	1 if yes, 0 if no
`ui`	uterine irritability	1 if yes, 0 if no
`lwd`	maternal weight before pregnancy	1 if weight $<$ 110 lb., 0 otherwise
`ptd`	history of premature labor	1 if yes, 0 if no
`c.age##lwd`	`age` main effects, `lwd` main effects, and their interaction	
`smoke##lwd`	`smoke` main effects, `lwd` main effects, and their interaction	

We first fit a model without the interaction terms by using `logit`.

```
. use https://www.stata-press.com/data/r18/lbw3
(Hosmer & Lemeshow data)
. logit low age lwd i.race smoke ptd ht ui
Iteration 0:  Log likelihood =   -117.336
Iteration 1:  Log likelihood =   -99.3982
Iteration 2:  Log likelihood = -98.780418
Iteration 3:  Log likelihood = -98.777998
Iteration 4:  Log likelihood = -98.777998
Logistic regression                                     Number of obs =    189
                                                        LR chi2(8)    =  37.12
                                                        Prob > chi2   = 0.0000
Log likelihood = -98.777998                             Pseudo R2     = 0.1582
```

low	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
age	-.0464796	.0373888	-1.24	0.214	-.1197603	.0268011
lwd	.8420615	.4055338	2.08	0.038	.0472299	1.636893
race						
Black	1.073456	.5150753	2.08	0.037	.0639273	2.082985
Other	.815367	.4452979	1.83	0.067	-.0574008	1.688135
smoke	.8071996	.404446	2.00	0.046	.0145001	1.599899
ptd	1.281678	.4621157	2.77	0.006	.3759478	2.187408
ht	1.435227	.6482699	2.21	0.027	.1646414	2.705813
ui	.6576256	.4666192	1.41	0.159	-.2569313	1.572182
_cons	-1.216781	.9556797	-1.27	0.203	-3.089878	.656317

To get the odds ratio for black smokers relative to white nonsmokers (the reference group), we type

```
. lincom 2.race + smoke, or
 ( 1)  [low]2.race + [low]smoke = 0
```

low	Odds ratio	Std. err.	z	P>|z|	[95% conf.	interval]
(1)	6.557805	4.744692	2.60	0.009	1.588176	27.07811

`lincom` computed $\exp(\beta_{\texttt{2.race}} + \beta_{\texttt{smoke}}) = 6.56$. To see the odds ratio for white smokers relative to black nonsmokers, we type

```
. lincom smoke - 2.race, or
 ( 1)  - [low]2.race + [low]smoke = 0
```

low	Odds ratio	Std. err.	z	P>|z|	[95% conf.	interval]
(1)	.7662425	.4430176	-0.46	0.645	.2467334	2.379603

Now let's add the interaction terms to the model (Hosmer and Lemeshow 1989, table 4.10). This time, we will use `logistic` rather than `logit`. By default, `logistic` displays odds ratios.

```
. logistic low i.race ht ui ptd c.age##lwd smoke##lwd
Logistic regression                                     Number of obs =    189
                                                        LR chi2(10)   =  42.66
                                                        Prob > chi2   = 0.0000
Log likelihood = -96.00616                              Pseudo R2     = 0.1818
```

low	Odds ratio	Std. err.	z	P>|z|	[95% conf.	interval]
race						
Black	2.95383	1.532789	2.09	0.037	1.068277	8.167465
Other	2.137589	.9919138	1.64	0.102	.8608708	5.307752
ht	3.893141	2.575201	2.05	0.040	1.064768	14.2346
ui	2.071284	.9931388	1.52	0.129	.8092926	5.301192
ptd	3.426633	1.615282	2.61	0.009	1.360252	8.632089
age	.9194513	.041896	-1.84	0.065	.8408967	1.005344
1.lwd	.1772934	.3312384	-0.93	0.354	.0045539	6.902367
lwd#c.age						
1	1.15883	.09602	1.78	0.075	.9851215	1.36317
smoke						
Smoker	3.168096	1.452378	2.52	0.012	1.289956	7.78076
smoke#lwd						
Smoker#1	.2447849	.2003996	-1.72	0.086	.0491956	1.217988
_cons	.599443	.6519163	-0.47	0.638	.0711271	5.051971

Note: **_cons** estimates baseline odds.

Hosmer and Lemeshow (1989, table 4.13) consider the effects of smoking (`smoke` = 1) and low maternal weight before pregnancy (`lwd` = 1). The effect of smoking among non–low-weight mothers (`lwd` = 0) is given by the odds ratio 3.17 for `smoke` in the `logistic` output. The effect of smoking among low-weight mothers is given by

```
. lincom 1.smoke + 1.smoke#1.lwd
 ( 1)  [low]1.smoke + [low]1.smoke#1.lwd = 0
```

low	Odds ratio	Std. err.	z	P>|z|	[95% conf.	interval]
(1)	.7755022	.574951	-0.34	0.732	.1813465	3.316323

We did not have to specify the `or` option. After `logistic`, `lincom` assumes `or` by default.

The effect of low weight (`lwd` = 1) is more complicated because we fit an `age` × `lwd` interaction. We must specify the age of mothers for the effect. The effect among 30-year-old nonsmokers is given by

```
. lincom 1.lwd + 30*1.lwd#c.age
 ( 1)  [low]1.lwd + 30*[low]1.lwd#c.age = 0
```

low	Odds ratio	Std. err.	z	P>|z|	[95% conf.	interval]
(1)	14.7669	13.5669	2.93	0.003	2.439264	89.39633

lincom computed $\exp(\beta_{\texttt{1wd}} + 30\beta_{\texttt{agelwd}}) = 14.8$. It may seem odd that we entered it as 1.lwd + 30*1.lwd#c.age, but remember that these terms are just lincom's (and test's) shorthands for _b[1.lwd] and _b[1.lwd#c.age]. We could have typed

```
. lincom _b[1.lwd] + 30*_b[1.lwd#c.age]
 ( 1)  [low]1.lwd + 30*[low]1.lwd#c.age = 0
```

low	Odds ratio	Std. err.	z	P>|z|	[95% conf.	interval]
(1)	14.7669	13.5669	2.93	0.003	2.439264	89.39633

◁

Multiple-equation models

lincom also works with multiple-equation models. The only difference is how you refer to the coefficients. Recall that for multiple-equation models, coefficients are referenced using the syntax

[*eqno*]*varname*

where *eqno* is the equation number or equation name and *varname* is the corresponding variable name for the coefficient; see [U] **13.5 Accessing coefficients and standard errors** and [R] **test** for details.

▷ Example 3

Let's consider example 4 from [R] **mlogit** (Tarlov et al. 1989; Wells et al. 1989).

```
. use https://www.stata-press.com/data/r18/sysdsn1
(Health insurance data)
. mlogit insure age male nonwhite i.site, nolog
Multinomial logistic regression                 Number of obs =    615
                                                LR chi2(10)   =  42.99
                                                Prob > chi2   = 0.0000
Log likelihood = -534.36165                     Pseudo R2     = 0.0387
```

insure	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
Indemnity	(base outcome)					
Prepaid						
age	-.011745	.0061946	-1.90	0.058	-.0238862	.0003962
male	.5616934	.2027465	2.77	0.006	.1643175	.9590693
nonwhite	.9747768	.2363213	4.12	0.000	.5115955	1.437958
site						
2	.1130359	.2101903	0.54	0.591	-.2989296	.5250013
3	-.5879879	.2279351	-2.58	0.010	-1.034733	-.1412433
_cons	.2697127	.3284422	0.82	0.412	-.3740222	.9134476
Uninsure						
age	-.0077961	.0114418	-0.68	0.496	-.0302217	.0146294
male	.4518496	.3674867	1.23	0.219	-.268411	1.17211
nonwhite	.2170589	.4256361	0.51	0.610	-.6171725	1.05129
site						
2	-1.211563	.4705127	-2.57	0.010	-2.133751	-.2893747
3	-.2078123	.3662926	-0.57	0.570	-.9257327	.510108
_cons	-1.286943	.5923219	-2.17	0.030	-2.447872	-.1260134

To see the estimate of the sum of the coefficient of `male` and the coefficient of `nonwhite` for the `Prepaid` outcome, we type

```
. lincom [Prepaid]male + [Prepaid]nonwhite
 ( 1)  [Prepaid]male + [Prepaid]nonwhite = 0
```

insure	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
(1)	1.53647	.3272489	4.70	0.000	.8950741	2.177866

To view the estimate as a ratio of relative risks (see [R] **mlogit** for the definition and interpretation), we specify the `rrr` option.

```
. lincom [Prepaid]male + [Prepaid]nonwhite, rrr
 ( 1)  [Prepaid]male + [Prepaid]nonwhite = 0
```

insure	RRR	Std. err.	z	P>|z|	[95% conf.	interval]
(1)	4.648154	1.521103	4.70	0.000	2.447517	8.827451

◁

Stored results

`lincom` stores the following in `r()`:

Scalars

`r(estimate)`	point estimate
`r(se)`	estimate of standard error
`r(df)`	degrees of freedom
`r(t)` or `r(z)`	t or z statistic
`r(p)`	p-value
`r(lb)`	lower bound of confidence interval
`r(ub)`	upper bound of confidence interval
`r(level)`	confidence level

References

Hosmer, D. W., Jr., and S. A. Lemeshow. 1989. *Applied Logistic Regression*. New York: Wiley.

Hosmer, D. W., Jr., S. A. Lemeshow, and R. X. Sturdivant. 2013. *Applied Logistic Regression*. 3rd ed. Hoboken, NJ: Wiley.

Tarlov, A. R., J. E. Ware, Jr., S. Greenfield, E. C. Nelson, E. Perrin, and M. Zubkoff. 1989. The medical outcomes study. An application of methods for monitoring the results of medical care. *Journal of the American Medical Association* 262: 925–930. https://doi.org/10.1001/jama.1989.03430070073033.

Wells, K. B., R. D. Hays, M. A. Burnam, W. H. Rogers, S. Greenfield, and J. E. Ware, Jr. 1989. Detection of depressive disorder for patients receiving prepaid or fee-for-service care. Results from the Medical Outcomes Survey. *Journal of the American Medical Association* 262: 3298–3302. https://doi.org/10.1001/jama.1989.03430230083030.

Also see

[R] **nlcom** — Nonlinear combinations of parameters

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

[R] **testnl** — Test nonlinear hypotheses after estimation

[SVY] **svy postestimation** — Postestimation tools for svy

[U] 13.5 Accessing coefficients and standard errors

[U] 20 Estimation and postestimation commands

Stata, Stata Press, and Mata are registered trademarks of StataCorp LLC. Stata and Stata Press are registered trademarks with the World Intellectual Property Organization of the United Nations. Other brand and product names are registered trademarks or trademarks of their respective companies. Copyright © 1985–2023 StataCorp LLC, College Station, TX, USA. All rights reserved.