Title

boxcox — Box–Cox regression models

Syntax

`boxcox` *depvar* [*indepvars*] [*if*] [*in*] [*weight*] [, *options*]

options	Description
Model	
`noconstant`	suppress constant term
`model(lhsonly)`	left-hand-side Box–Cox model; the default
`model(rhsonly)`	right-hand-side Box–Cox model
`model(lambda)`	both sides Box–Cox model with same parameter
`model(theta)`	both sides Box–Cox model with different parameters
`notrans(`*varlist*`)`	nontransformed independent variables
Reporting	
`level(#)`	set confidence level; default is `level(95)`
`lrtest`	perform likelihood-ratio test
Maximization	
`nolog`	suppress full-model iteration log
`nologlr`	suppress restricted-model `lrtest` iteration log
maximize_options	control the maximization process; seldom used

depvar and *indepvars* may contain time-series operators; see [U] **11.4.4 Time-series varlists**.

`bootstrap`, `by`, `jackknife`, `rolling`, `statsby`, and `xi` are allowed; see [U] **11.1.10 Prefix commands**.

Weights are not allowed with the `bootstrap` prefix; see [R] **bootstrap**.

`fweight`s and `iweight`s are allowed; see [U] **11.1.6 weight**.

See [U] **20 Estimation and postestimation commands** for more capabilities of estimation commands.

Menu

Statistics > Linear models and related > Box-Cox regression

Description

`boxcox` finds the maximum likelihood estimates of the parameters of the Box–Cox transform, the coefficients on the independent variables, and the standard deviation of the normally distributed errors for a model in which *depvar* is regressed on *indepvars*. You can fit the following models:

Option	Estimates
`lhsonly`	$y_j^{(\theta)} = \beta_1 x_{1j} + \beta_2 x_{2j} + \cdots + \beta_k x_{kj} + \epsilon_j$
`rhsonly`	$y_j \;= \beta_1 x_{1j}^{(\lambda)} + \beta_2 x_{2j}^{(\lambda)} + \cdots + \beta_k x_{kj}^{(\lambda)} + \epsilon_j$
`rhsonly notrans()`	$y_j \;= \beta_1 x_{1j}^{(\lambda)} + \beta_2 x_{2j}^{(\lambda)} + \cdots + \beta_k x_{kj}^{(\lambda)} + \gamma_1 z_{1j} + \cdots + \gamma_l z_{lj} + \epsilon_j$
`lambda`	$y_j^{(\lambda)} = \beta_1 x_{1j}^{(\lambda)} + \beta_2 x_{2j}^{(\lambda)} + \cdots + \beta_k x_{kj}^{(\lambda)} + \epsilon_j$
`lambda notrans()`	$y_j^{(\lambda)} = \beta_1 x_{1j}^{(\lambda)} + \beta_2 x_{2j}^{(\lambda)} + \cdots + \beta_k x_{kj}^{(\lambda)} + \gamma_1 z_{1j} + \cdots + \gamma_l z_{lj} + \epsilon_j$
`theta`	$y_j^{(\theta)} = \beta_1 x_{1j}^{(\lambda)} + \beta_2 x_{2j}^{(\lambda)} + \cdots + \beta_k x_{kj}^{(\lambda)} + \epsilon_j$
`theta notrans()`	$y_j^{(\theta)} = \beta_1 x_{1j}^{(\lambda)} + \beta_2 x_{2j}^{(\lambda)} + \cdots + \beta_k x_{kj}^{(\lambda)} + \gamma_1 z_{1j} + \cdots + \gamma_l z_{lj} + \epsilon_j$

Any variable to be transformed must be strictly positive.

Options

Model

`noconstant`; see [R] **estimation options**.

`model( lhsonly | rhsonly | lambda | theta )` specifies which of the four models to fit.

`model(lhsonly)` applies the Box–Cox transform to *depvar* only. `model(lhsonly)` is the default.

`model(rhsonly)` applies the transform to the *indepvars* only.

`model(lambda)` applies the transform to both *depvar* and *indepvars*, and they are transformed by the same parameter.

`model(theta)` applies the transform to both *depvar* and *indepvars*, but this time, each side is transformed by a separate parameter.

`notrans(`*varlist*`)` specifies that the variables in *varlist* be included as nontransformed independent variables.

Reporting

`level(`*#*`)`; see [R] **estimation options**.

`lrtest` specifies that a likelihood-ratio test of significance be performed and reported for each independent variable.

Maximization

`nolog` suppresses the iteration log when fitting the full model.

`nologlr` suppresses the iteration log when fitting the restricted models required by the `lrtest` option.

maximize_options: `iterate(`*#*`)` and `from(`*init_specs*`)`; see [R] **maximize**.

Model	Initial value specification
`lhsonly`	`from(`θ_0`, copy)`
`rhsonly`	`from(`λ_0`, copy)`
`lambda`	`from(`λ_0`, copy)`
`theta`	`from(`λ_0 θ_0`, copy)`

Remarks and examples

Remarks are presented under the following headings:

Introduction

The Box–Cox transform

$$y^{(\lambda)} = \frac{y^\lambda - 1}{\lambda}$$

has been widely used in applied data analysis. Box and Cox (1964) developed the transformation and argued that the transformation could make the residuals more closely normal and less heteroskedastic. Cook and Weisberg (1982) discuss the transform in this light. Because the transform embeds several popular functional forms, it has received some attention as a method for testing functional forms, in particular,

$$y^{(\lambda)} = \begin{cases} y - 1 & \text{if } \lambda = 1 \\ \ln(y) & \text{if } \lambda = 0 \\ 1 - 1/y & \text{if } \lambda = -1 \end{cases}$$

Davidson and MacKinnon (1993) discuss this use of the transform. Atkinson (1985) also gives a good general treatment.

Theta model

`boxcox` obtains the maximum likelihood estimates of the parameters for four different models. The most general of the models, the `theta` model, is

$$y_j^{(\theta)} = \beta_0 + \beta_1 x_{1j}^{(\lambda)} + \beta_2 x_{2j}^{(\lambda)} + \cdots + \beta_k x_{kj}^{(\lambda)} + \gamma_1 z_{1j} + \gamma_2 z_{2j} + \cdots + \gamma_l z_{lj} + \epsilon_j$$

where $\epsilon \sim N(0, \sigma^2)$. Here the dependent variable, y, is subject to a Box–Cox transform with parameter θ. Each of the *indepvars*, $x_1, x_2, \ldots, x_k$, is transformed by a Box–Cox transform with parameter λ. The $z_1, z_2, \ldots, z_l$ specified in the `notrans()` option are independent variables that are not transformed.

Box and Cox (1964) argued that this transformation would leave behind residuals that more closely follow a normal distribution than those produced by a simple linear regression model. Bear in mind that the normality of ϵ is assumed and that `boxcox` obtains maximum likelihood estimates of the $k + l + 4$ parameters under this assumption. `boxcox` does not choose λ and θ so that the residuals are approximately normally distributed. If you are interested in this type of transformation to normality, see the official Stata commands `lnskew0` and `bcskew0` in [R] **lnskew0**. However, those commands work on a more restrictive model in which none of the independent variables is transformed.

▷ Example 1

Below we fit a `theta` model to a nonrepresentative extract of the Second National Health and Nutrition Examination Survey (NHANES II) dataset discussed in McDowell et al. (1981).

We model individual-level diastolic blood pressure (`bpdiast`) as a function of the transformed variables body mass index (`bmi`) and cholesterol level (`tcresult`) and of the untransformed variables age (`age`) and sex (`sex`).

```
. use http://www.stata-press.com/data/r13/nhanes2
. boxcox bpdiast bmi tcresult, notrans(age sex) model(theta) lrtest
Fitting comparison model
Iteration 0:   log likelihood =  -41178.61
Iteration 1:   log likelihood =  -41032.51
Iteration 2:   log likelihood = -41032.488
Iteration 3:   log likelihood = -41032.488
Fitting full model
Iteration 0:   log likelihood = -39928.606
Iteration 1:   log likelihood = -39775.026
Iteration 2:   log likelihood = -39774.987
Iteration 3:   log likelihood = -39774.987
Fitting comparison models for LR tests
Iteration 0:   log likelihood = -39947.144
Iteration 1:   log likelihood =  -39934.55
Iteration 2:   log likelihood = -39934.516
Iteration 3:   log likelihood = -39934.516
Iteration 0:   log likelihood =  -39906.96
Iteration 1:   log likelihood =  -39896.63
Iteration 2:   log likelihood = -39896.629
Iteration 0:   log likelihood = -40464.599
Iteration 1:   log likelihood = -40459.752
Iteration 2:   log likelihood = -40459.604
Iteration 3:   log likelihood = -40459.604
Iteration 0:   log likelihood = -39829.859
Iteration 1:   log likelihood = -39815.576
Iteration 2:   log likelihood = -39815.575
                                                  Number of obs   =      10351
                                                  LR chi2(5)      =    2515.00
Log likelihood = -39774.987                       Prob > chi2     =      0.000
```

bpdiast	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
/lambda	.6383286	.1577601	4.05	0.000	.3291245	.9475327
/theta	.1988197	.0454088	4.38	0.000	.1098201	.2878193

Estimates of scale-variant parameters

	Coef.	chi2(df)	P>chi2(df)	df of chi2
Notrans				
age	.003811	319.060	0.000	1
sex	-.1054887	243.284	0.000	1
_cons	5.835555			
Trans				
bmi	.0872041	1369.235	0.000	1
tcresult	.004734	81.177	0.000	1
/sigma	.3348267			

Test H0:	Restricted log likelihood	chi2	Prob > chi2
theta=lambda = -1	-40162.898	775.82	0.000
theta=lambda = 0	-39790.945	31.92	0.000
theta=lambda = 1	-39928.606	307.24	0.000

The output is composed of the iteration logs and three distinct tables. The first table contains a standard header for a maximum likelihood estimator and a standard output table for the Box–Cox transform parameters. The second table contains the estimates of the scale-variant parameters. The third table contains the output from likelihood-ratio tests on three standard functional form specifications.

The right-hand-side and the left-hand-side transformations each add to the regression fit at the 1% significance level and are both positive but less than 1. All the variables have significant impacts on diastolic blood pressure, `bpdiast`. As expected, the transformed variables—the body mass index, `bmi`, and cholesterol level, `tcresult`—contribute to higher blood pressure. The last output table shows that the linear, multiplicative inverse, and log specifications are strongly rejected. ◁

❑ Technical note

Spitzer (1984) showed that the Wald tests of the joint significance of the coefficients of the right-hand-side variables, either transformed or untransformed, are not invariant to changes in the scale of the transformed dependent variable. Davidson and MacKinnon (1993) also discuss this point. This problem demonstrates that Wald statistics can be manipulated in nonlinear models. Lafontaine and White (1986) analyze this problem numerically, and Phillips and Park (1988) analyze it by using Edgeworth expansions. See Drukker (2000b) for a more detailed discussion of this issue. Because the parameter estimates and their Wald tests are not scale invariant, no Wald tests or confidence intervals are reported for these parameters. However, when the `lrtest` option is specified, likelihood-ratio tests are performed and reported. Schlesselman (1971) showed that, if a constant is included in the model, the parameter estimates of the Box–Cox transforms are scale invariant. For this reason, we strongly recommend that you not use the `noconstant` option.

The `lrtest` option does not perform a likelihood-ratio test on the constant, so no value for this statistic is reported. Unless the data are properly scaled, the restricted model does not often converge. For this reason, no likelihood-ratio test on the constant is performed by the `lrtest` option. However, if you have a special interest in performing this test, you can do so by fitting the constrained model separately. If problems with convergence are encountered, rescaling the data by their means may help. ❑

Lambda model

A less general model than the one above is called the `lambda` model. It specifies that the same parameter be used in both the left-hand-side and right-hand-side transformations. Specifically,

$$y_j^{(\lambda)} = \beta_0 + \beta_1 x_{1j}^{(\lambda)} + \beta_2 x_{2j}^{(\lambda)} + \cdots + \beta_k x_{kj}^{(\lambda)} + \gamma_1 z_{1j} + \gamma_2 z_{2j} + \cdots + \gamma_l z_{lj} + \epsilon_j$$

where $\epsilon \sim N(0, \sigma^2)$. Here the *depvar* variable, y, and each of the *indepvars*, $x_1, x_2, \ldots, x_k$, is transformed by a Box–Cox transform with the common parameter λ. Again the $z_1, z_2, \ldots, z_l$ are independent variables that are not transformed.

Left-hand-side-only model

Even more restrictive than a common transformation parameter is transforming the dependent variable only. Because the dependent variable is on the left-hand side of the equation, this model is known as the `lhsonly` model. Here you are estimating the parameters of the model

$$y_j^{(\theta)} = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \cdots + \beta_k x_{kj} + \epsilon_j$$

where $\epsilon \sim N(0, \sigma^2)$. Here only the *depvar*, y, is transformed by a Box–Cox transform with the parameter θ.

▷ Example 2

In this example, we model the transform of diastolic blood pressure as a linear combination of the untransformed body mass index, cholesterol level, age, and sex.

```
. boxcox bpdiast bmi tcresult age sex, model(lhsonly) lrtest nolog nologlr
Fitting comparison model
Fitting full model
Fitting comparison models for LR tests
                                                 Number of obs   =      10351
                                                 LR chi2(4)      =    2509.56
Log likelihood = -39777.709                      Prob > chi2     =      0.000

------------------------------------------------------------------------------
     bpdiast |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      /theta |   .2073268   .0452895     4.58   0.000     .1185611    .2960926
------------------------------------------------------------------------------

Estimates of scale-variant parameters
----------------------------------------------------------
             |      Coef.   chi2(df)  P>chi2(df)  df of chi2
-------------+--------------------------------------------
Notrans      |
         bmi |   .0272628   1375.841    0.000         1
    tcresult |   .0006929     82.380    0.000         1
         age |   .0040141    334.117    0.000         1
         sex |  -.1122274    263.219    0.000         1
       _cons |   6.302855
-------------+--------------------------------------------
      /sigma |   .3476615
----------------------------------------------------------

---------------------------------------------------------
   Test         Restricted      LR statistic     P-value
   H0:        log likelihood        chi2       Prob > chi2
---------------------------------------------------------
theta = -1      -40146.678         737.94         0.000
theta =  0      -39788.241          21.06         0.000
theta =  1      -39928.606         301.79         0.000
---------------------------------------------------------
```

The maximum likelihood estimate of the transformation parameter for this model is positive and significant. Once again, all the scale-variant parameters are significant, and we find a positive impact of body mass index (`bmi`) and cholesterol levels (`tcresult`) on the transformed diastolic blood pressure (`bpdiast`). This model rejects the linear, multiplicative inverse, and log specifications. ◁

Right-hand-side-only model

The fourth model leaves the *depvar* alone and transforms a subset of the *indepvars* using the parameter λ. This is the `rhsonly` model. In this model, the *depvar*, y, is given by

$$y_j = \beta_0 + \beta_1 x_{1j}^{(\lambda)} + \beta_2 x_{2j}^{(\lambda)} + \cdots + \beta_k x_{kj}^{(\lambda)} + \gamma_1 z_{1j} + \gamma_2 z_{2j} + \cdots + \gamma_l z_{lj} + \epsilon_j$$

where $\epsilon \sim N(0, \sigma^2)$. Here each of the *indepvars*, $x_1, x_2, \ldots, x_k$, is transformed by a Box–Cox transform with the parameter λ. Again the $z_1, z_2, \ldots, z_l$ are independent variables that are not transformed.

▷ Example 3

Now we consider a rhsonly model in which the regressors sex and age are not transformed.

```
. boxcox bpdiast bmi tcresult, notrans(sex age) model(rhsonly) lrtest nolog
> nologlr
Fitting full model
Fitting comparison models for LR tests
                                                   Number of obs   =      10351
                                                   LR chi2(5)      =    2500.79
Log likelihood = -39928.212                        Prob > chi2     =      0.000

------------------------------------------------------------------------------
     bpdiast |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     /lambda |   .8658841   .1522387     5.69   0.000     .5675018    1.164266
------------------------------------------------------------------------------

Estimates of scale-variant parameters
----------------------------------------------------------
             |      Coef.  chi2(df)  P>chi2(df)  df of chi2
-------------+--------------------------------------------
Notrans      |
         sex |  -3.544042   235.020    0.000          1
         age |    .128809   311.754    0.000          1
       _cons |   50.01498
-------------+--------------------------------------------
Trans        |
         bmi |   1.418215  1396.709   0.000           1
    tcresult |   .0462964    78.500   0.000           1
-------------+--------------------------------------------
      /sigma |    11.4557
----------------------------------------------------------

---------------------------------------------------------
   Test          Restricted      LR statistic      P-value
   H0:         log likelihood        chi2       Prob > chi2
---------------------------------------------------------
lambda = -1       -39989.331        122.24          0.000
lambda =  0       -39942.945         29.47          0.000
lambda =  1       -39928.606          0.79          0.375
---------------------------------------------------------
```

The maximum likelihood estimate of the transformation parameter in this model is positive and significant at the 1% level. The transformed bmi coefficient behaves as expected, and the remaining scale-variant parameters are significant at the 1% level. This model rejects the multiplicative inverse and log specifications strongly. However, we cannot reject the hypothesis that the model is linear.

◁

Stored results

boxcox stores the following in e():

Scalars	
e(N)	number of observations
e(ll)	log likelihood
e(chi2)	LR statistic of full vs. comparison
e(df_m)	full model degrees of freedom
e(ll0)	log likelihood of the restricted model
e(df_r)	restricted model degrees of freedom
e(ll_t1)	log likelihood of model $\lambda = \theta = 1$
e(chi2_t1)	LR of $\lambda = \theta = 1$ vs. full model
e(p_t1)	p-value of $\lambda = \theta = 1$ vs. full model
e(ll_tm1)	log likelihood of model $\lambda = \theta = -1$
e(chi2_tm1)	LR of $\lambda = \theta = -1$ vs. full model
e(p_tm1)	p-value of $\lambda = \theta = -1$ vs. full model
e(ll_t0)	log likelihood of model $\lambda = \theta = 0$
e(chi2_t0)	LR of $\lambda = \theta = 0$ vs. full model
e(p_t0)	p-value of $\lambda = \theta = 0$ vs. full model
e(rank)	rank of e(V)
e(ic)	number of iterations
e(rc)	return code
Macros	
e(cmd)	boxcox
e(cmdline)	command as typed
e(depvar)	name of dependent variable
e(model)	lhsonly, rhsonly, lambda, or theta
e(wtype)	weight type
e(wexp)	weight expression
e(ntrans)	yes if nontransformed *indepvars*
e(chi2type)	LR; type of model χ^2 test
e(lrtest)	lrtest, if requested
e(properties)	b V
e(predict)	program used to implement predict
e(marginsnotok)	predictions disallowed by margins
Matrices	
e(b)	coefficient vector
e(V)	variance–covariance matrix of the estimators (see note below)
e(pm)	p-values for LR tests on *indepvars*
e(df)	degrees of freedom of LR tests on *indepvars*
e(chi2m)	LR statistics for tests on *indepvars*
Functions	
e(sample)	marks estimation sample

e(V) contains all zeros, except for the elements that correspond to the parameters of the Box–Cox transform.

Methods and formulas

In the internal computations,

$$y^{(\lambda)} = \begin{cases} \frac{y^{\lambda}-1}{\lambda} & \text{if } |\lambda| > 10^{-10} \\ \ln(y) & \text{otherwise} \end{cases}$$

The unconcentrated log likelihood for the `theta` model is

$$\ln L = \left(\frac{-N}{2}\right)\left\{\ln(2\pi) + \ln(\sigma^2)\right\} + (\theta - 1)\sum_{i=1}^{N} \ln(y_i) - \left(\frac{1}{2\sigma^2}\right)\text{SSR}$$

where

$$\text{SSR} = \sum_{i=1}^{N}(y_i^{(\theta)} - \beta_0 + \beta_1 x_{i1}^{(\lambda)} + \beta_2 x_{i2}^{(\lambda)} + \cdots + \beta_k x_{ik}^{(\lambda)} + \gamma_1 z_{i1} + \gamma_2 z_{i2} + \cdots + \gamma_l z_{il})^2$$

Writing the SSR in matrix form,

$$\text{SSR} = (\mathbf{y}^{(\theta)} - \mathbf{X}^{(\lambda)}\mathbf{b}' - \mathbf{Z}\mathbf{g}')'(\mathbf{y}^{(\theta)} - \mathbf{X}^{(\lambda)}\mathbf{b}' - \mathbf{Z}\mathbf{g}')$$

where $\mathbf{y}^{(\theta)}$ is an $N \times 1$ vector of elementwise transformed data, $\mathbf{X}^{(\lambda)}$ is an $N \times k$ matrix of elementwise transformed data, $\mathbf{Z}$ is an $N \times l$ matrix of untransformed data, $\mathbf{b}$ is a $1 \times k$ vector of coefficients, and $\mathbf{g}$ is a $1 \times l$ vector of coefficients. Letting

$$\mathbf{W}_\lambda = \left(\mathbf{X}^{(\lambda)} \ \mathbf{Z}\right)$$

be the horizontal concatenation of $\mathbf{X}^{(\lambda)}$ and $\mathbf{Z}$ and

$$\mathbf{d}' = \begin{pmatrix} \mathbf{b}' \\ \mathbf{g}' \end{pmatrix}$$

be the vertical concatenation of the coefficients yields

$$\text{SSR} = (\mathbf{y}^{(\theta)} - \mathbf{W}_\lambda \mathbf{d}')'(\mathbf{y}^{(\theta)} - \mathbf{W}_\lambda \mathbf{d}')$$

For given values of λ and θ, the solutions for $\mathbf{d}'$ and σ^2 are

$$\widehat{\mathbf{d}}' = (W_\lambda' W_\lambda)^{-1} W_\lambda' y^{(\theta)}$$

and

$$\widehat{\sigma}^{\,2} = \frac{1}{N}\left(\mathbf{y}^{(\theta)} - W_\lambda \widehat{d'}\right)' \left(y^{(\theta)} - W_\lambda \widehat{d'}\right)$$

Substituting these solutions into the log-likelihood function yields the concentrated log-likelihood function

$$\ln L_c = \left(-\frac{N}{2}\right)\left\{\ln(2\pi) + 1 + \ln(\widehat{\sigma}^{\,2})\right\} + (\theta - 1)\sum_{i=1}^{N} \ln(y_i)$$

Similar calculations yield the concentrated log-likelihood function for the `lambda` model,

$$\ln L_c = \left(-\frac{N}{2}\right)\left\{\ln(2\pi)+1+\ln(\widehat{\sigma}^{\,2})\right\}+(\lambda-1)\sum_{i=1}^{N}\ln(y_i)$$

the `lhsonly` model,

$$\ln L_c = \left(-\frac{N}{2}\right)\left\{\ln(2\pi)+1+\ln(\widehat{\sigma}^{\,2})\right\}+(\theta-1)\sum_{i=1}^{N}\ln(y_i)$$

and the `rhsonly` model,

$$\ln L_c = \left(-\frac{N}{2}\right)\left\{\ln(2\pi)+1+\ln(\widehat{\sigma}^{\,2})\right\}$$

where $\widehat{\sigma}^{\,2}$ is specific to each model and is defined analogously to that in the `theta` model.

References

Atkinson, A. C. 1985. *Plots, Transformations, and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis*. Oxford: Oxford University Press.

Box, G. E. P., and D. R. Cox. 1964. An analysis of transformations. *Journal of the Royal Statistical Society, Series B* 26: 211–252.

Carroll, R. J., and D. Ruppert. 1988. *Transformation and Weighting in Regression*. New York: Chapman & Hall.

Cook, R. D., and S. Weisberg. 1982. *Residuals and Influence in Regression*. New York: Chapman & Hall/CRC.

Davidson, R., and J. G. MacKinnon. 1993. *Estimation and Inference in Econometrics*. New York: Oxford University Press.

Drukker, D. M. 2000a. sg130: Box–Cox regression models. *Stata Technical Bulletin* 54: 27–36. Reprinted in *Stata Technical Bulletin Reprints*, vol. 9, pp. 307–319. College Station, TX: Stata Press.

——. 2000b. sg131: On the manipulability of Wald tests in Box–Cox regression models. *Stata Technical Bulletin* 54: 36–42. Reprinted in *Stata Technical Bulletin Reprints*, vol. 9, pp. 319–327. College Station, TX: Stata Press.

Lafontaine, F., and K. J. White. 1986. Obtaining any Wald statistic you want. *Economics Letters* 21: 35–40.

Lindsey, C., and S. J. Sheather. 2010a. Power transformation via multivariate Box–Cox. *Stata Journal* 10: 69–81.

——. 2010b. Optimal power transformation via inverse response plots. *Stata Journal* 10: 200–214.

McDowell, A., A. Engel, J. T. Massey, and K. Maurer. 1981. Plan and operation of the Second National Health and Nutrition Examination Survey, 1976–1980. *Vital and Health Statistics* 1(15): 1–144.

Phillips, P. C. B., and J. Y. Park. 1988. On the formulation of Wald tests of nonlinear restrictions. *Econometrica* 56: 1065–1083.

Schlesselman, J. J. 1971. Power families: A note on the Box and Cox transformation. *Journal of the Royal Statistical Society, Series B* 33: 307–311.

Spitzer, J. J. 1984. Variance estimates in models with the Box–Cox transformation: Implications for estimation and hypothesis testing. *Review of Economics and Statistics* 66: 645–652.

Also see

[R] **boxcox postestimation** — Postestimation tools for boxcox

[R] **lnskew0** — Find zero-skewness log or Box–Cox transform

[R] **regress** — Linear regression

[U] **20 Estimation and postestimation commands**