logit — Logistic regression, reporting coefficients										
Description	Quick start	Menu	Syntax	Options						
Remarks and examples	Stored results	Methods and formulas	References	Also see						

Description

logit fits a logit model for a binary response by maximum likelihood; it models the probability of a positive outcome given a set of regressors. *depvar* equal to nonzero and nonmissing (typically *depvar* equal to one) indicates a positive outcome, whereas *depvar* equal to zero indicates a negative outcome.

Also see [R] logistic; logistic displays estimates as odds ratios. Many users prefer the logistic command to logit. Results are the same regardless of which you use—both are the maximum-likelihood estimator. Several auxiliary commands that can be run after logit, probit, or logistic estimation are described in [R] logistic postestimation.

Quick start

Logit model of y on x1 and x2

logit y x1 x2

Add indicators for categorical variable a

logit y x1 x2 i.a

With cluster-robust standard errors for clustering by levels of cvar

logit y x1 x2 i.a, vce(cluster cvar)

Save separate coefficient estimates for each level of cvar to myresults.dta statsby _b, by(cvar) saving(myresults): logit y x1 x2 i.a

Adjust for complex survey design using svyset data svy: logit y x1 x2 i.a

Menu

 $Statistics > Binary \ outcomes > Logistic \ regression$

Syntax

logit depvar [indepvars] [if] [in] [weight] [, options]

options	Description
Model	
<u>nocons</u> tant <u>off</u> set(<i>varname</i>) asis <u>const</u> raints(<i>constraints</i>)	suppress constant term include <i>varname</i> in model with coefficient constrained to 1 retain perfect predictor variables apply specified linear constraints
SE/Robust	
vce(<i>vcetype</i>)	<i>vcetype</i> may be oim, opg, <u>r</u> obust, <u>cl</u> uster <i>clustvar</i> , <u>boot</u> strap, or <u>jackknife</u>
Reporting	
<u>l</u> evel(#)	set confidence level; default is level(95)
or	report odds ratios
<u>nocnsr</u> eport	do not display constraints
display_options	control columns and column formats, row spacing, line width, display of omitted variables and base and empty cells, and factor-variable labeling
Maximization	
maximize_options	control the maximization process; seldom used
nocoef	do not display coefficient table; seldom used
<u>col</u> linear	keep collinear variables
<u>coefl</u> egend	display legend instead of statistics

indepvars may contain factor variables; see [U] 11.4.3 Factor variables.

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

bayes, bayesboot, bootstrap, by, collect, fmm, fp, jackknife, mfp, mi estimate, nestreg, rolling, statsby, stepwise, and svy are allowed; see [U] 11.1.10 Prefix commands. For more details, see [BAYES] bayes: logit and [FMM] fmm: logit.

vce(bootstrap) and vce(jackknife) are not allowed with the mi estimate prefix; see [MI] mi estimate.

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

vce(), nocoef, and weights are not allowed with the svy prefix; see [SVY] svy.

fweights, iweights, and pweights are allowed; see [U] 11.1.6 weight.

nocoef, collinear, and coeflegend do not appear in the dialog box.

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

Options

Model

noconstant, offset(varname), constraints(constraints); see [R] Estimation options.

asis forces retention of perfect predictor variables and their associated perfectly predicted observations and may produce instabilities in maximization; see [R] probit.

SE/Robust

vce(vcetype) specifies the type of standard error reported, which includes types that are derived from asymptotic theory (oim, opg), that are robust to some kinds of misspecification (robust), that allow for intragroup correlation (cluster *clustvar*), and that use bootstrap or jackknife methods (bootstrap, jackknife); see [R] vce_option.

Reporting

level(#); see [R] Estimation options.

or reports the estimated coefficients transformed to odds ratios, that is, e^b rather than b. Standard errors and confidence intervals are similarly transformed. This option affects how results are displayed, not how they are estimated. or may be specified at estimation or when replaying previously estimated results.

nocnsreport; see [R] Estimation options.

```
display_options: noci, nopvalues, noomitted, vsquish, noemptycells, baselevels,
allbaselevels, nofvlabel, fvwrap(#), fvwrapon(style), cformat(%fmt), pformat(%fmt),
sformat(%fmt), and nolstretch; see [R] Estimation options.
```

Maximization

maximize_options: difficult, technique(algorithm_spec), iterate(#), [no]log, trace, gradient, showstep, hessian, showtolerance, tolerance(#), ltolerance(#), nrtolerance(#), nonrtolerance, and from(init_specs); see [R] Maximize. These options are seldom used.

The following options are available with logit but are not shown in the dialog box:

nocoef specifies that the coefficient table not be displayed. This option is sometimes used by program writers but is of no use interactively.

collinear, coeflegend; see [R] Estimation options.

Remarks and examples

Remarks are presented under the following headings:

Basic usage Model identification

Basic usage

logit fits maximum likelihood models with dichotomous dependent (left-hand-side) variables coded as 0/1 (or, more precisely, coded as 0 and not-0).

For grouped data or data in binomial form, a probit model can be fit using glm with the family(binomial) and link(logit) options.

Example 1

We have data on the make, weight, and mileage rating of 22 foreign and 52 domestic automobiles. We wish to fit a logit model explaining whether a car is foreign on the basis of its weight and mileage. Here is an overview of our data:

```
. use https://www.stata-press.com/data/r19/auto
(1978 automobile data)
. keep make mpg weight foreign
. describe
Contains data from https://www.stata-press.com/data/r19/auto.dta
                            74
                                                 1978 automobile data
 Observations:
                             4
    Variables:
                                                 13 Apr 2024 17:45
                                                 ( dta has notes)
Variable
               Storage
                         Display
                                     Value
                          format
                                     label
                                                 Variable label
    name
                  type
make
                 str18
                         %-18s
                                                 Make and model
                         %8.0g
                                                 Mileage (mpg)
                 int
mpg
weight
                 int
                         %8.0gc
                                                 Weight (lbs.)
foreign
                 byte
                         %8.0g
                                     origin
                                                 Car origin
Sorted by: foreign
     Note: Dataset has changed since last saved.
. inspect foreign
foreign: Car origin
                                                   Number of observations
                                                Total
                                                            Integers
                                                                        Nonintegers
   #
                               Negative
   #
                                                   52
                                                                  52
                               Zero
   #
                               Positive
                                                   22
                                                                  22
   #
   #
       #
                                                   74
                                                                  74
                               Total
   #
       #
                               Missing
                                                   74
Ó
                       1
   (2 unique values)
```

foreign is labeled and all values are documented in the label.

The variable foreign takes on two unique values, 0 and 1. The value 0 denotes a domestic car, and 1 denotes a foreign car.

The model that we wish to fit is

 $\Pr(\texttt{foreign} = 1) = F(\beta_0 + \beta_1 \texttt{weight} + \beta_2 \texttt{mpg})$

where $F(z) = e^{z}/(1 + e^{z})$ is the cumulative logistic distribution.

To fit this model, we type

. logit foreig	gn weight mpg					
Iteration 0: Iteration 1: Iteration 2: Iteration 3: Iteration 4: Iteration 5:	Log likelihoo Log likelihoo Log likelihoo Log likelihoo Log likelihoo Log likelihoo	d = -45.03 $d = -29.238$ $d = -27.244$ $d = -27.175$ $d = -27.175$ $d = -27.175$	321 536 139 277 156 156			
Logistic regre	ession 1 = -27.175156				Number of ob: LR chi2(2) Prob > chi2 Pseudo R2	s = 74 = 35.72 = 0.0000 = 0.3966
foreign	Coefficient	Std. err.	Z	P> z	[95% conf.	interval]
weight mpg _cons	0039067 1685869 13.70837	.0010116 .0919175 4.518709	-3.86 -1.83 3.03	0.000 0.067 0.002	0058894 3487418 4.851859	001924 .011568 22.56487

We find that heavier cars are less likely to be foreign and that cars yielding better gas mileage are also less likely to be foreign, at least holding the weight of the car constant.

Technical note

Stata interprets a value of 0 as a negative outcome (failure) and treats all other values (except missing) as positive outcomes (successes). Thus if your dependent variable takes on the values 0 and 1, then 0 is interpreted as failure and 1 as success. If your dependent variable takes on the values 0, 1, and 2, then 0 is still interpreted as failure, but both 1 and 2 are treated as successes.

If you prefer a more formal mathematical statement, when you type logit y x, Stata fits the model

$$\Pr(y_j \neq 0 \mid \mathbf{x}_j) = \frac{\exp(\mathbf{x}_j \boldsymbol{\beta})}{1 + \exp(\mathbf{x}_j \boldsymbol{\beta})}$$

Model identification

The logit command has one more feature, and it is probably the most useful. logit automatically checks the model for identification and, if it is underidentified, drops whatever variables and observations are necessary for estimation to proceed. (logistic, probit, and ivprobit do this as well.)

Example 2

Have you ever fit a logit model where one or more of your independent variables perfectly predicted one or the other outcome?

For instance, consider the following data:

Outcome y	Independent variable x
0	1
0	1
0	0
1	0

Say that we wish to predict the outcome on the basis of the independent variable. The outcome is always zero whenever the independent variable is one. In our data, Pr(y = 0 | x = 1) = 1, which means that the logit coefficient on x must be minus infinity with a corresponding infinite standard error. At this point, you may suspect that we have a problem.

Unfortunately, not all such problems are so easily detected, especially if you have a lot of independent variables in your model. If you have ever had such difficulties, you have experienced one of the more unpleasant aspects of computer optimization. The computer has no idea that it is trying to solve for an infinite coefficient as it begins its iterative process. All it knows is that at each step, making the coefficient a little bigger, or a little smaller, works wonders. It continues on its merry way until either 1) the whole thing comes crashing to the ground when a numerical overflow error occurs or 2) it reaches some predetermined cutoff that stops the process. In the meantime, you have been waiting. The estimates that you finally receive, if you receive any at all, may be nothing more than numerical roundoff.

Stata watches for these sorts of problems, alerts us, fixes them, and properly fits the model.

Let's return to our automobile data. Among the variables we have in the data is one called repair, which takes on three values. A value of 1 indicates that the car has a poor repair record, 2 indicates an average record, and 3 indicates a better-than-average record. Here is a tabulation of our data:

```
. use https://www.stata-press.com/data/r19/repair, clear
(1978 automobile data)
. tabulate foreign repair
                            Repair
Car origin
                                   2
                                               3
                                                        Total
                       1
  Domestic
                      10
                                  27
                                               9
                                                           46
   Foreign
                       0
                                   3
                                               9
                                                           12
     Total
                      10
                                  30
                                              18
                                                           58
```

All the cars with poor repair records (repair = 1) are domestic. If we were to attempt to predict foreign on the basis of the repair records, the predicted probability for the repair = 1 category would have to be zero. This in turn means that the logit coefficient must be minus infinity, and that would set most computer programs buzzing.

4

Let's try Stata on this problem.

. logit foreig	gn b3.repair					
note: 1.repain 1.repain	r != 0 predict r omitted and	s failure po 10 obs not n	erfectly; used.			
Iteration 0: Iteration 1: Iteration 2: Iteration 3: Iteration 4:	Log likelihoo Log likelihoo Log likelihoo Log likelihoo Log likelihoo	d = -26.9920 $d = -22.483$ $d = -22.2300$ $d = -22.229$ $d = -22.229$	087 187 498 139 138			
Logistic regre Log likelihood	ession 1 = -22.229138	3			Number of oh LR chi2(1) Prob > chi2 Pseudo R2	bs = 48 = 9.53 = 0.0020 = 0.1765
foreign	Coefficient	Std. err.	z	P> z	[95% conf.	interval]
repair 1 2	0 -2.197225	(empty) .7698003	-2.85	0.004	-3.706005	6884436
_cons	-1.85e-17	.4714045	-0.00	1.000	9239359	.9239359

Remember that all the cars with poor repair records (repair = 1) are domestic, so the model cannot be fit, or at least it cannot be fit if we restrict ourselves to finite coefficients. Stata noted that fact "note: 1.repair !=0 predicts failure perfectly". This is Stata's mathematically precise way of saying what we said in English. When repair is 1, the car is domestic.

Stata then went on to say "1.repair omitted and 10 obs not used". This is Stata eliminating the problem. First 1.repair had to be removed from the model because it would have an infinite coefficient. Then, the 10 observations that led to the problem had to be eliminated, as well, so as not to bias the remaining coefficients in the model. The 10 observations that are not used are the 10 domestic cars that have poor repair records.

Stata then fit what was left of the model, using the remaining observations. Because no observations remained for cars with poor repair records, Stata reports "(empty)" in the row for repair = 1.

Technical note

Stata is pretty smart about catching problems like this. It will catch "one-way causation by a dummy variable", as we demonstrated above.

Stata also watches for "two-way causation", that is, a variable that perfectly determines the outcome, both successes and failures. Here Stata says, "so-and-so predicts outcome perfectly" and stops. Statistics dictates that no model can be fit.

Stata also checks your data for collinear variables; it will say, "so-and-so omitted because of collinearity". No observations need to be eliminated in this case, and model fitting will proceed without the offending variable. It will also catch a subtle problem that can arise with continuous data. For instance, if we were estimating the chances of surviving the first year after an operation, and if we included in our model age, and if all the persons over 65 died within the year, Stata would say, "age > 65 predicts failure perfectly". It would then inform us about the fix-up it takes and fit what can be fit of our model.

logit (and logistic, probit, and ivprobit) will also occasionally display messages such as

Note: 4 failures and 0 successes completely determined.

There are two causes for a message like this. The first—and most unlikely—case occurs when a continuous variable (or a combination of a continuous variable with other continuous or dummy variables) is simply a great predictor of the dependent variable. Consider Stata's auto.dta dataset with 6 observations removed.

. use https:// (1978 automobi	/www.stata-pre ile data)	ss.com/data	/r19/auto)		
. drop if fore (6 observation	eign==0 & gear ns deleted)	_ratio > 3.	1			
. logit foreig	gn mpg weight	gear_ratio,	nolog			
Logistic regre Log likelihood	Number of ob LR chi2(3) Prob > chi2 Pseudo R2	s = 68 = 72.64 = 0.0000 = 0.8484				
foreign	Coefficient	Std. err.	z	P> z	[95% conf.	interval]
mpg weight gear_ratio _cons	4944907 0060919 15.70509 -21.39527	.2655508 .003101 8.166234 25.41486	-1.86 -1.96 1.92 -0.84	0.063 0.049 0.054 0.400	-1.014961 0121698 300436 -71.20747	.0259792 000014 31.71061 28.41694

Note: 4 failures and 0 successes completely determined.

There are no missing standard errors in the output. If you receive the "completely determined" message and have one or more missing standard errors in your output, see the second case discussed below.

Note gear_ratio's large coefficient. logit thought that the 4 observations with the smallest predicted probabilities were essentially predicted perfectly.

If this happens to you, you do not have to do anything. Computationally, the model is sound. The second case discussed below requires careful examination.

The second case occurs when the independent terms are all dummy variables or continuous ones with repeated values (for example, age). Here one or more of the estimated coefficients will have missing standard errors. For example, consider this dataset consisting of 6 observations.

```
. use https://www.stata-press.com/data/r19/logitxmpl, clear
```

```
. list, separator(0)
```

	у	x1	x2
1.	0	0	0
2.	0	0	0
з.	0	1	0
4.	1	1	0
5.	0	0	1
6.	1	0	1

. logit y x1 x2

Iteration	0:	Log	likelihood	1 =	-3.81908	5					
Iteration	1:	Log	likelihood	1 = -	2.952733	6					
Iteration	2:	Log	likelihood	1 = -	2.811028	2					
Iteration	3:	Log	likelihood	1 = -	2.781197	3					
Iteration	4:	Log	likelihood	1 = -	2.774610	7					
Iteration	5:	Log	likelihood	1 = -	2.773012	8					
(output on	nitted)									
Iteration	296:	Log	g likelihoo	od =	-2.77258	87 (n	lot concar	ve)			
Iteration	297 :	: Log	g likelihoo	od =	-2.77258	87 (n	ot conca	ve)			
Iteration	298:	Log	g likelihoo	od =	-2.77258	87 (n	lot concar	ve)			
Iteration	299:	Log	g likelihoo	od =	-2.77258	87 (n	lot concar	ve)			
Iteration	300:	: Log	g likelihoo	od =	-2.77258	87 (n	lot concar	ve)			
convergen	ce no	ot ac	chieved								
Logistic 1	regre	essio	on					Number o	f obs	3 =	6
0	0							LR chi2(1)	=	2.09
								Prob > c	hi2	= 0	. 1480
Log likel:	ihood	d = -	-2.7725887					Pseudo R	.2	= 0	.2740
	у	Coe	efficient	Std.	err.	z	P> z	[95% c	onf.	inte	rval]
	x1		18.3704		2	9.19	0.000	14.450	47	22.3	29033
	x2		18.3704								

Note: 2 failures and 0 successes completely determined. convergence not achieved r(430);

1.414214

-18.3704

_cons

Three things are happening here. First, logit iterates almost forever and then declares nonconvergence. Second, logit can fit the outcome (y = 0) for the covariate pattern x1 = 0 and x2 = 0 (that is, the first two observations) perfectly. This observation is the "2 failures and 0 successes completely determined". Third, if this observation is excluded, then x1, x2, and the constant are collinear.

-12.99

0.000

-21.14221

-15.5986

This is the cause of the nonconvergence, the message "completely determined", and the missing standard errors. It happens when you have a covariate pattern (or patterns) with only one outcome and there is collinearity when the observations corresponding to this covariate pattern are excluded. If this happens to you, confirm the causes. First, identify the covariate pattern with only one outcome. (For your data, replace x1 and x2 with the independent variables of your model.)

```
. egen pattern = group(x1 x2)
. quietly logit y x1 x2, iterate(100)
convergence not achieved
. predict p
(option pr assumed; Pr(y))
. summarize p
    Variable
                       Obs
                                  Mean
                                           Std. dev.
                                                            Min
                                                                       Max
           р
                         6
                              .3333333
                                           .2581989
                                                      1.05e-08
                                                                         .5
```

If successes were completely determined, that means that there are predicted probabilities that are almost 1. If failures were completely determined, that means that there are predicted probabilities that are almost 0. The latter is the case here, so we locate the corresponding value of pattern:

•	tabulate pa	ttern if p < 1e-7								
	group(x1 x2)	Freq.	Percent	Cum.						
	1	2	100.00	100.00						
	Total	2	100.00							

Once we omit this covariate pattern from the estimation sample, logit can deal with the collinearity:

. logit y x1 x note: x2 omitt	x2 if pattern ted because of	!= 1, nolog f collineari	ty.			
Logistic regre	ession				Number of ob LR chi2(1) Prob > chi2	s = 4 = 0.00 = 1.0000
Log likelihood	1 = -2.7725887	7			Pseudo R2	= 0.0000
У	Coefficient	Std. err.	z	P> z	[95% conf.	interval]
x1 x2	0	2 (omitted)	0.00	1.000	-3.919928	3.919928
_cons	0	1.414214	0.00	1.000	-2.771808	2.771808

. logit y x1,	nolog					
Logistic regre	ession				Number of ob	s = 6
с с					LR chi2(1)	= 0.37
					Prob > chi2	= 0.5447
Log likelihood	d = -3.6356349)			Pseudo R2	= 0.0480
У	Coefficient	Std. err.	z	P> z	[95% conf.	interval]
x1	1.098612	1.825742	0.60	0.547	-2.479776	4.677001
_cons	-1.098612	1.154701	-0.95	0.341	-3.361784	1.164559

We omit the collinear variable. Then we must decide whether to include or omit the observations with pattern = 1. We could include them,

or exclude them,

. logit y x1 i	if pattern !=	1, nolog				
Logistic regre	Number of ob LR chi2(1) Prob > chi2	s = 4 = 0.00 = 1.0000				
Log likelihood	Pseudo R2	= 0.0000				
у	Coefficient	Std. err.	Z	P> z	[95% conf.	interval]
x1 _cons	0 0	2 1.414214	0.00	1.000	-3.919928 -2.771808	3.919928 2.771808

If the covariate pattern that predicts outcome perfectly is meaningful, you may want to exclude these observations from the model. Here you would report that covariate pattern such and such predicted outcome perfectly and that the best model for the rest of the data is But, more likely, the perfect prediction was simply the result of having too many predictors in the model. Then you would omit the extraneous variables from further consideration and report the best model for all the data.

Stored results

logit stores the following in e():

1415	
e(N)	number of observations
e(N_cds)	number of completely determined successes
e(N_cdf)	number of completely determined failures
e(k)	number of parameters
e(k_eq)	number of equations in e(b)
e(k_eq_model)	number of equations in overall model test
e(k_dv)	number of dependent variables
e(df_m)	model degrees of freedom
e(r2_p)	pseudo- R^2
e(11)	log likelihood
e(11_0)	log likelihood, constant-only model
e(N_clust)	number of clusters
e(chi2)	χ^2
e(p)	<i>p</i> -value for model test
e(rank)	rank of e(V)
e(ic)	number of iterations
e(rc)	return code
e(converged)	1 if converged, 0 otherwise

Macros	
e(cmd)	logit
e(cmdline)	command as typed
e(depvar)	name of dependent variable
e(wtype)	weight type
e(wexp)	weight expression
e(title)	title in estimation output
e(clustvar)	name of cluster variable
e(offset)	linear offset variable
e(chi2type)	Wald or LR; type of model χ^2 test
e(vce)	vcetype specified in vce()
e(vcetype)	title used to label Std. err.
e(opt)	type of optimization
e(which)	max or min; whether optimizer is to perform maximization or minimization
e(ml_method)	type of ml method
e(user)	name of likelihood-evaluator program
e(technique)	maximization technique
e(properties)	b V
e(estat_cmd)	program used to implement estat
e(predict)	program used to implement predict
e(marginsok)	predictions allowed by margins
e(marginsnotok)	predictions disallowed by margins
e(asbalanced)	factor variables fvset as asbalanced
e(asobserved)	factor variables fvset as asobserved
Matrices	
e(b)	coefficient vector
e(Cns)	constraints matrix
e(ilog)	iteration log (up to 20 iterations)
e(gradient)	gradient vector
e(mns)	vector of means of the independent variables
e(rules)	information about perfect predictors
e(V)	variance-covariance matrix of the estimators
e(V_modelbased)	model-based variance
Functions	
e(sample)	marks estimation sample
1 I I I I I I I I I I I I I I I I I I I	1

In addition to the above, the following is stored in r():

Matrices	
r(table)	

matrix containing the coefficients with their standard errors, test statistics, *p*-values, and confidence intervals

Note that results stored in r() are updated when the command is replayed and will be replaced when any r-class command is run after the estimation command.

Methods and formulas

Cramer (2003, chap. 9) surveys the prehistory and history of the logit model. The word "logit" was coined by Berkson (1944) and is analogous to the word "probit". For an introduction to probit and logit, see, for example, Aldrich and Nelson (1984), Cameron and Trivedi (2022), Jones (2007), Long (1997), Long and Freese (2014), Pampel (2021), or Powers and Xie (2008).

The likelihood function for logit is

$$\mathrm{ln}L = \sum_{j \in S} w_j \ln F(\mathbf{x}_j \mathbf{b}) + \sum_{j \notin S} w_j \ln \{1 - F(\mathbf{x}_j \mathbf{b})\}$$

where S is the set of all observations j, such that $y_j \neq 0$, $F(z) = e^z/(1 + e^z)$, and w_j denotes the optional weights. $\ln L$ is maximized as described in [R] Maximize.

This command supports the Huber/White/sandwich estimator of the variance and its clustered version using vce(robust) and vce(cluster *clustvar*), respectively. See [P] **_robust**, particularly *Maximum likelihood estimators* and *Methods and formulas*. The scores are calculated as $\mathbf{u}_j = \{1 - F(\mathbf{x}_j \mathbf{b})\}\mathbf{x}_j$ for the positive outcomes and $-F(\mathbf{x}_j \mathbf{b})\mathbf{x}_j$ for the negative outcomes.

logit also supports estimation with survey data. For details on VCEs with survey data, see [SVY] Variance estimation.

Joseph Berkson (1899–1982) was born in New York City and studied at the College of the City of New York, Columbia, and Johns Hopkins, earning both an MD and a doctorate in statistics. He then worked at Johns Hopkins before moving to the Mayo Clinic in 1931 as a biostatistician. Among many other contributions, his most influential one drew upon a long-sustained interest in the logistic function, especially his 1944 paper on bioassay, in which he introduced the term "logit". Berkson was a frequent participant in controversy—sometimes humorous, sometimes bitter—on subjects such as the evidence for links between smoking and various diseases and the relative merits of probit and logit methods and of different calculation methods.

References

- Aldrich, J. H., and F. D. Nelson. 1984. Linear Probability, Logit, and Probit Models. Newbury Park, CA: Sage. https://doi.org/10.4135/9781412984744.
- Archer, K. J., and S. A. Lemeshow. 2006. Goodness-of-fit test for a logistic regression model fitted using survey sample data. Stata Journal 6: 97–105.
- Berkson, J. 1944. Application of the logistic function to bio-assay. *Journal of the American Statistical Association* 39: 357–365. https://doi.org/10.2307/2280041.
- Buis, M. L. 2010a. Direct and indirect effects in a logit model. Stata Journal 10: 11-29.

_____. 2010b. Stata tip 87: Interpretation of interactions in nonlinear models. Stata Journal 10: 305–308.

- Cameron, A. C., and P. K. Trivedi. 2022. Microeconometrics Using Stata. 2nd ed. College Station, TX: Stata Press.
- Corral, P., and M. Terbish. 2015. Generalized maximum entropy estimation of discrete choice models. *Stata Journal* 15: 512–522.
- Cramer, J. S. 2003. Logit Models from Economics and Other Fields. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511615412.
- Drukker, D. M. 2016. Probability differences and odds ratios measure conditional-on-covariate effects and populationparameter effects. The Stata Blog: Not Elsewhere Classified. https://blog.stata.com/2016/07/26/probability-differencesand-odds-ratios-measure-conditional-on-covariate-effects-and-population-parameter-effects/.
- Fernandez-Felix, B. M., E. García-Esquinas, A. Muriel, A. Royuela, and J. Zamora. 2021. Bootstrap internal validation command for predictive logistic regression models. *Stata Journal* 21: 498–509.
- Hilbe, J. M. 2009. Logistic Regression Models. Boca Raton, FL: Chapman and Hall/CRC.
- Hosmer, D. W., Jr., S. A. Lemeshow, and R. X. Sturdivant. 2013. Applied Logistic Regression. 3rd ed. Hoboken, NJ: Wiley.
- Jones, A. M. 2007. Applied Econometrics for Health Economists: A Practical Guide. 2nd ed. Abingdon, UK: Radcliffe.
- Judge, G. G., W. E. Griffiths, R. C. Hill, H. Lütkepohl, and T.-C. Lee. 1985. The Theory and Practice of Econometrics. 2nd ed. New York: Wiley.
- Long, J. S. 1997. Regression Models for Categorical and Limited Dependent Variables. Thousand Oaks, CA: Sage.

- Long, J. S., and J. Freese. 2014. Regression Models for Categorical Dependent Variables Using Stata. 3rd ed. College Station, TX: Stata Press.
- Miranda, A., and S. Rabe-Hesketh. 2006. Maximum likelihood estimation of endogenous switching and sample selection models for binary, ordinal, and count variables. *Stata Journal* 6: 285–308.
- Mitchell, M. N., and X. Chen. 2005. Visualizing main effects and interactions for binary logit models. *Stata Journal* 5: 64–82.
- O'Fallon, W. M. 2005. "Berkson, Joseph". In Encyclopedia of Biostatistics, edited by P. Armitage and T. Colton, vol. 1: 376–381. Chichester, UK: Wiley. https://doi.org/10.1002/0470011815.b2a17006.
- Orsini, N., R. Bellocco, and P. C. Sjölander. 2013. Doubly robust estimation in generalized linear models. *Stata Journal* 13: 185–205.
- Pampel, F. C. 2021. Logistic Regression: A Primer. 2nd ed. Thousand Oaks, CA: Sage.
- Pedace, R. 2013. Econometrics for Dummies. Hoboken, NJ: Wiley.
- Pollock, P. H., III, and B. C. Edwards. 2019. A Stata Companion to Political Analysis. 4th ed. Thousand Oaks, CA: CQ Press.
- Powers, D. A., and Y. Xie. 2008. Statistical Methods for Categorical Data Analysis. 2nd ed. Bingley, UK: Emerald.
- Pregibon, D. 1981. Logistic regression diagnostics. Annals of Statistics 9: 705–724. https://doi.org/10.1214/aos/ 1176345513.
- Uberti, L. J. 2022. Interpreting logit models. Stata Journal 22: 60-76.
- Xu, J., and J. S. Long. 2005. Confidence intervals for predicted outcomes in regression models for categorical outcomes. Stata Journal 5: 537–559.

Also see

- [R] logit postestimation Postestimation tools for logit
- [R] brier Brier score decomposition
- [R] **cloglog** Complementary log–log regression
- [R] exlogistic Exact logistic regression
- [R] logistic Logistic regression, reporting odds ratios
- [R] npregress kernel Nonparametric kernel regression
- [R] npregress series Nonparametric series regression
- [R] **probit** Probit regression
- [R] roc Receiver operating characteristic (ROC) analysis
- [R] ziologit Zero-inflated ordered logit regression
- [BAYES] bayes: logit Bayesian logistic regression, reporting coefficients
- [FMM] fmm: logit Finite mixtures of logistic regression models
- [LASSO] Lasso intro Introduction to lasso
- [ME] melogit Multilevel mixed-effects logistic regression
- [MI] Estimation Estimation commands for use with mi estimate
- [PSS-2] power logistic general Power analysis for logistic regression: General case⁺
- [PSS-2] power logistic onebin Power analysis for logistic regression with one binary covariate⁺
- [PSS-2] power logistic twobin Power analysis for logistic regression with two binary covariates⁺
- [SVY] svy estimation Estimation commands for survey data
- [XT] xtlogit Fixed-effects, random-effects, and population-averaged logit models
- [U] 20 Estimation and postestimation commands

Stata, Stata Press, Mata, NetCourse, and NetCourseNow are registered trademarks of StataCorp LLC. Stata and Stata Press are registered trademarks with the World Intellectual Property Organization of the United Nations. StataNow is a trademark of StataCorp LLC. Other brand and product names are registered trademarks or trademarks of their respective companies. Copyright © 1985–2025 StataCorp LLC, College Station, TX, USA. All rights reserved.



For suggested citations, see the FAQ on citing Stata documentation.