margins, contrast — Contrasts of margins

Description

`margins` with the `contrast` option or with contrast operators performs contrasts of margins. This extends the capabilities of `contrast` to any of the nonlinear responses, predictive margins, or other margins that can be estimated by `margins`.

Quick start

Reference category contrasts of the predictive margins for `a` after `logit y a##b x1`

```
margins r.a
```

Contrasts of predictive margins for `a` with the previous level

```
margins ar.a
```

Test the equality of predictive margins for `a`

```
margins a, contrast
```

Reference category contrasts of predictive margins for `x1` = 10 and `x1` = 20 with `x1` = 0

```
margins, at(x1=(0 10 20)) contrast(atcontrast(r))
```

Average partial effect of increasing `x1` by 100 for each observation after `probit y x1 x2`

```
margins, at((asobserved) _all)  at(x1=generate(x1+100))    ///
     contrast(atcontrast(r))
```

Menu

Statistics > Postestimation

Syntax

margins [*marginlist*] [*if*] [*in*] [*weight*] [, contrast *margins_options*]

margins [*marginlist*] [*if*] [*in*] [*weight*] [, contrast(*suboptions*) *margins_options*]

where *marginlist* is a list of factor variables or interactions that appear in the current estimation results. The variables may be typed with or without contrast operators, and you may use any factor-variable syntax:

```
. margins sex##group, contrast
. margins sex##g.group, contrast
. margins sex@group, contrast
```

See the operators (*op.*) table in [R] **contrast** for the list of contrast operators. Contrast operators may also be specified on the variables in margins's over() and within() options to perform contrasts across the levels of those variables.

suboptions	Description
Contrast	
overall	add a joint hypothesis test for all specified contrasts
lincom	treat user-defined contrasts as linear combinations
predict(*op*[._predict])	apply the *op*. contrast operator to the groups defined by multiple predict() options
atcontrast(*op*[._at])	apply the *op*. contrast operator to the groups defined by at()
predictjoint	test jointly across all groups defined by multiple predict() options
atjoint	test jointly across all groups defined by at()
overjoint	test jointly across all levels of the unoperated over() variables
withinjoint	test jointly across all levels of the unoperated within() variables
marginswithin	perform contrasts within the levels of the unoperated terms in *marginlist*
cieffects	show effects table with confidence intervals
pveffects	show effects table with p-values
effects	show effects table with confidence intervals and p-values
nowald	suppress table of Wald tests
noatlevels	report only the overall Wald test for terms that use the within @ or nested | operator
nosvyadjust	compute unadjusted Wald tests for survey results

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

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

Suboptions

Contrast

overall specifies that a joint hypothesis test over all terms be performed.

`lincom` specifies that user-defined contrasts be treated as linear combinations. The default is to require that all user-defined contrasts sum to zero. (Summing to zero is part of the definition of a contrast.)

`predict(`*op*[`._predict`]`)` specifies that the *op*`.` contrast operator be applied to the groups defined by multiple specifications of `margins`'s `predict()` option. The default behavior, by comparison, is to perform tests and contrasts within these groups.

`atcontrast(`*op*[`._at`]`)` specifies that the *op*`.` contrast operator be applied to the groups defined by the `at()` option(s). The default behavior, by comparison, is to perform tests and contrasts within the groups defined by the `at()` option(s).

See example 6 in *Remarks and examples*.

`predictjoint` specifies that joint tests be performed across all groups defined by multiple specifications of `margins`'s `predict()` option. The default behavior, by comparison, is to perform contrasts and tests within each group.

`atjoint` specifies that joint tests be performed across all groups defined by the `at()` option. The default behavior, by comparison, is to perform contrasts and tests within each group.

See example 5 in *Remarks and examples*.

`overjoint` specifies how unoperated variables in the `over()` option are treated.

Each variable in the `over()` option may be specified either with or without a contrast operator. For contrast-operated variables, the specified contrast comparisons are always performed.

`overjoint` specifies that joint tests be performed across all levels of the unoperated variables. The default behavior, by comparison, is to perform contrasts and tests within each combination of levels of the unoperated variables.

See example 3 in *Remarks and examples*.

`withinjoint` specifies how unoperated variables in the `within()` option are treated.

Each variable in the `within()` option may be specified either with or without a contrast operator. For contrast-operated variables, the specified contrast comparisons are always performed.

`withinjoint` specifies that joint tests be performed across all levels of the unoperated variables. The default behavior, by comparison, is to perform contrasts and tests within each combination of levels of the unoperated variables.

`marginswithin` specifies how unoperated variables in *marginlist* are treated.

Each variable in *marginlist* may be specified either with or without a contrast operator. For contrast-operated variables, the specified contrast comparisons are always performed.

`marginswithin` specifies that contrasts and tests be performed within each combination of levels of the unoperated variables. The default behavior, by comparison, is to perform joint tests across all levels of the unoperated variables.

See example 4 in *Remarks and examples*.

`cieffects` specifies that a table containing a confidence interval for each individual contrast be reported.

`pveffects` specifies that a table containing a p-value for each individual contrast be reported.

`effects` specifies that a single table containing a confidence interval and p-value for each individual contrast be reported.

`nowald` suppresses the table of Wald tests.

`noatlevels` indicates that only the overall Wald test be reported for each term containing within or nested (`@` or `|`) operators.

`nosvyadjust` is for use with `svy` estimation commands. It specifies that the Wald test be carried out without the default adjustment for the design degrees of freedom. That is to say the test is carried out as $W/k \sim F(k,d)$ rather than as $(d-k+1)W/(kd) \sim F(k,d-k+1)$, where k is the dimension of the test and d is the total number of sampled PSUs minus the total number of strata.

Remarks and examples

stata.com

Remarks are presented under the following headings:

Contrasts of margins

▷ Example 1

Estimating contrasts of margins is as easy as adding a contrast operator to the variable name. Let's review *Example 2: A simple case after logistic* of [R] **margins**. Variable `sex` is coded `0` for males and `1` for females.

```
. use https://www.stata-press.com/data/r18/margex
(Artificial data for margins)

. logistic outcome i.sex i.group
  (output omitted)

. margins sex

Predictive margins                                      Number of obs = 3,000
Model VCE: OIM

Expression: Pr(outcome), predict()
```

	Margin	Delta-method std. err.	z	P>|z|	[95% conf. interval]	
sex						
Male	.1286796	.0111424	11.55	0.000	.106841	.1505182
Female	.1905087	.0089719	21.23	0.000	.1729241	.2080933

The first margin, 0.13, is the average probability of a positive outcome, treating everyone as if they were male. The second margin, 0.19, is the average probability of a positive outcome, treating everyone as if they were female. We can compare females with males by rerunning `margins` and adding a contrast operator:

```
. margins r.sex
Contrasts of predictive margins                              Number of obs = 3,000
Model VCE: OIM
Expression: Pr(outcome), predict()
```

	df	chi2	P>chi2
sex	1	16.61	0.0000

	Contrast	Delta-method std. err.	[95% conf.	interval]
sex (Female vs Male)	.0618291	.0151719	.0320927	.0915656

The `r.` prefix for `sex` is the reference-category contrast operator—see [R] **contrast**. (The default reference category is zero, the lowest value of `sex`.) Contrast operators in a *marginlist* work just as they do in the *termlist* of a `contrast` command.

The contrast estimate of 0.06 says that unconditional on group, females on average are about 6% more likely than males to have a positive outcome. The χ^2 statistic of 16.61 shows that the contrast is significantly different from zero.

You may be surprised that we did not need to include the `contrast` option to estimate our contrast. If we had included the option, our output would not have changed:

```
. margins r.sex, contrast
Contrasts of predictive margins                              Number of obs = 3,000
Model VCE: OIM
Expression: Pr(outcome), predict()
```

	df	chi2	P>chi2
sex	1	16.61	0.0000

	Contrast	Delta-method std. err.	[95% conf.	interval]
sex (Female vs Male)	.0618291	.0151719	.0320927	.0915656

The `contrast` option is useful mostly for its suboptions, which control the output and how contrasts are estimated in more complicated situations. But `contrast` may be specified on its own (without contrast operators or suboptions) if we do not need estimates or confidence intervals:

```
. margins sex group, contrast
Contrasts of predictive margins                              Number of obs = 3,000
Model VCE: OIM
Expression: Pr(outcome), predict()
```

	df	chi2	P>chi2
sex	1	16.61	0.0000
group	2	225.76	0.0000

Each χ^2 statistic is a joint test of constituent contrasts. The test for `group` has two degrees of freedom because `group` has three levels.

◁

Contrasts and the over() option

▷ Example 2

It is common to estimate margins at combinations of factor levels, and `margins, contrast` includes several suboptions for contrasting such margins. Let's fit a model with two categorical predictors and their interaction:

```
. logistic outcome agegroup##group

Logistic regression                                     Number of obs =  3,000
                                                        LR chi2(8)    = 520.64
                                                        Prob > chi2   = 0.0000
Log likelihood = -1105.7504                             Pseudo R2     = 0.1906

------------------------------------------------------------------------------
     outcome | Odds ratio   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
    agegroup |
      30-39  |    3.54191   2.226951     2.01   0.044     1.032882    12.14576
        40+  |   16.23351   9.611879     4.71   0.000     5.086452    51.80955
             |
       group |
          2  |    .834507   .5663738    -0.27   0.790     .2206611     3.15598
          3  |   .2146739   .1772903    -1.86   0.062      .042541    1.083306
             |
   agegroup# |
       group |
    30-39#2  |   .4426927   .3358505    -1.07   0.283     .1000772    1.958257
    30-39#3  |    1.16088   1.103521     0.16   0.875     .1801538    7.480508
      40+#2  |    .440672   .3049393    -1.18   0.236     .1135259     1.71055
      40+#3  |   .4407892   .4034666    -0.89   0.371     .0732998    2.650693
             |
       _cons |   .0379747   .0223371    -5.56   0.000     .0119897    .1202762
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.
```

Each of `agegroup` and `group` has three levels. To compare each age group with the reference category on the probability scale, we can again use `margins` with the `r.` contrast operator.

```
. margins r.agegroup

Contrasts of predictive margins                             Number of obs = 3,000
Model VCE: OIM

Expression: Pr(outcome), predict()

---------------------------------------------------------
                      |         df        chi2     P>chi2
----------------------+----------------------------------
             agegroup |
     (30-39 vs 20-29) |          1       10.04     0.0015
       (40+ vs 20-29) |          1      224.44     0.0000
                Joint |          2      238.21     0.0000
---------------------------------------------------------

------------------------------------------------------------------
                  |            Delta-method
                  |   Contrast   std. err.     [95% conf. interval]
------------------+-----------------------------------------------
         agegroup |
 (30-39 vs 20-29) |    .044498    .0140448      .0169706    .0720253
   (40+ vs 20-29) |   .2059281    .0137455      .1789874    .2328688
------------------------------------------------------------------
```

Our model includes an interaction, though, so it would be nice to estimate the contrasts separately for each value of `group`. We need the `over()` option:

```
. margins r.agegroup, over(group)

Contrasts of predictive margins                             Number of obs = 3,000
Model VCE: OIM

Expression: Pr(outcome), predict()
Over:       group

-----------------------------------------------------------
                        |         df        chi2     P>chi2
------------------------+----------------------------------
         agegroup@group |
     (30-39 vs 20-29) 1 |          1        6.94     0.0084
     (30-39 vs 20-29) 2 |          1        1.18     0.2783
     (30-39 vs 20-29) 3 |          1        3.10     0.0783
       (40+ vs 20-29) 1 |          1      173.42     0.0000
       (40+ vs 20-29) 2 |          1       57.77     0.0000
       (40+ vs 20-29) 3 |          1        5.12     0.0236
                  Joint |          6      266.84     0.0000
-----------------------------------------------------------

--------------------------------------------------------------------
                    |            Delta-method
                    |   Contrast   std. err.     [95% conf. interval]
--------------------+-----------------------------------------------
     agegroup@group |
 (30-39 vs 20-29) 1 |   .0819713    .0311208      .0209757     .142967
 (30-39 vs 20-29) 2 |   .0166206    .0153309     -.0134275    .0466686
 (30-39 vs 20-29) 3 |   .0243462    .0138291     -.0027584    .0514507
   (40+ vs 20-29) 1 |   .3447797    .0261811      .2934658    .3960937
   (40+ vs 20-29) 2 |   .1540882    .0202722      .1143554     .193821
   (40+ vs 20-29) 3 |   .0470319    .0207774       .006309    .0877548
--------------------------------------------------------------------
```

The effect of `agegroup` appears to be greatest for the first level of `group`.

Including a variable in the `over()` option is not equivalent to including the variable in the main *marginlist*. The variables in the *marginlist* are manipulated in the analysis, so that we can measure, for example, the effect of being in age group 3 and not age group 1. (The manipulation could be mimicked by running `replace` and then `predict`, but the manipulations actually performed by `margins` do not change the data in memory.) The variables in the `over()` option are not so manipulated—the values of the `over()` variables are left as they were observed, and the *marginlist* variables are manipulated separately for each observed `over()` group.

◁

The overjoint suboption

▷ Example 3

Each variable in an `over()` option may be specified with or without contrast operators. Our option `over(group)` did not include a contrast operator, so `margins` estimated the contrasts separately for each level of `group`. If we had instead specified `over(r.group)`, we would have received differences of the contrasts:

```
. margins r.agegroup, over(r.group)

Contrasts of predictive margins                        Number of obs = 3,000
Model VCE: OIM

Expression: Pr(outcome), predict()
Over:       group

------------------------------------------------------------------
                               |         df        chi2     P>chi2
-------------------------------+----------------------------------
                group#agegroup |
    (2 vs 1) (30-39 vs 20-29)  |          1        3.55     0.0596
      (2 vs 1) (40+ vs 20-29)  |          1       33.17     0.0000
    (3 vs 1) (30-39 vs 20-29)  |          1        2.86     0.0906
      (3 vs 1) (40+ vs 20-29)  |          1       79.36     0.0000
                        Joint  |          4       83.88     0.0000
------------------------------------------------------------------

--------------------------------------------------------------------------------
                           |            Delta-method
                           |   Contrast   std. err.     [95% conf. interval]
---------------------------+----------------------------------------------------
            group#agegroup |
(2 vs 1) (30-39 vs 20-29)  |  -.0653508   .0346921      -.133346    .0026445
  (2 vs 1) (40+ vs 20-29)  |  -.1906915   .0331121       -.25559   -.1257931
(3 vs 1) (30-39 vs 20-29)  |  -.0576252   .0340551     -.1243719    .0091216
  (3 vs 1) (40+ vs 20-29)  |  -.2977479   .0334237     -.3632572   -.2322386
--------------------------------------------------------------------------------
```

The contrasts are double differences: the estimate of −0.19, for example, says that the difference in the probability of success between age group 3 and age group 1 is smaller in group 2 than in group 1. We can jointly test pairs of the double differences with the `overjoint` suboption:

```
. margins r.agegroup, over(group) contrast(overjoint)

Contrasts of predictive margins                        Number of obs = 3,000
Model VCE: OIM

Expression: Pr(outcome), predict()
Over:       group

-----------------------------------------------------------------
                              |         df        chi2     P>chi2
------------------------------+----------------------------------
               group#agegroup |
    (joint) (30-39 vs 20-29)  |          2        3.62     0.1641
      (joint) (40+ vs 20-29)  |          2       79.45     0.0000
                       Joint  |          4       83.88     0.0000
-----------------------------------------------------------------
```

The `contrast(overjoint)` option overrides the default behavior of `over()` and requests joint tests over the levels of the unoperated variable `group`. The χ^2 statistic of 3.62 tests that the first and third contrasts from the previous table are jointly zero. The χ^2 statistic of 79.45 jointly tests the other pair of contrasts.

◁

The marginswithin suboption

▷ Example 4

Another suboption that may usefully be combined with `over()` is `marginswithin`. `marginswithin` requests that contrasts be performed within the levels of unoperated variables in the main *marginlist*, instead of performing them jointly across the levels. `marginswithin` affects only unoperated variables because contrast operators take precedence over suboptions.

Let's first look at the default behavior, which occurs when `marginswithin` is not specified:

```
. margins agegroup, over(r.group) contrast(effects)
Contrasts of predictive margins                         Number of obs = 3,000
Model VCE: OIM

Expression: Pr(outcome), predict()
Over:       group

─────────────────────────────────────────────────
                  │         df        chi2     P>chi2
──────────────────┼──────────────────────────────
   group#agegroup │
(2 vs 1) (joint)  │          2       33.94     0.0000
(3 vs 1) (joint)  │          2       83.38     0.0000
            Joint │          4       83.88     0.0000
─────────────────────────────────────────────────

──────────────────────────────────────────────────────────────────────────
             │            Delta-method
             │   Contrast   std. err.      z    P>|z|     [95% conf. interval]
─────────────┼────────────────────────────────────────────────────────────
      group# │
   agegroup  │
  (2 vs 1)   │
    (30-39   │
         vs  │
      base)  │  -.0653508   .0346921   -1.88   0.060     -.133346    .0026445
  (2 vs 1)   │
      (40+   │
         vs  │
      base)  │  -.1906915   .0331121   -5.76   0.000      -.25559   -.1257931
  (3 vs 1)   │
    (30-39   │
         vs  │
      base)  │  -.0576252   .0340551   -1.69   0.091    -.1243719    .0091216
  (3 vs 1)   │
      (40+   │
         vs  │
      base)  │  -.2977479   .0334237   -8.91   0.000    -.3632572   -.2322386
──────────────────────────────────────────────────────────────────────────
```

Here `agegroup` in the main *marginlist* is an unoperated variable, so `margins` by default performs joint tests across the levels of `agegroup`: the χ^2 statistic of 33.94, for example, jointly tests whether the first two contrast estimates in the lower table differ significantly from zero.

When we specify `marginswithin`, the contrasts will instead be performed within the levels of `agegroup`:

```
. margins agegroup, over(r.group) contrast(marginswithin effects)

Contrasts of predictive margins                        Number of obs = 3,000
Model VCE: OIM

Expression: Pr(outcome), predict()
Over:       group
```

	df	chi2	P>chi2
group@agegroup			
(2 vs 1) 20-29	1	0.06	0.7991
(2 vs 1) 30-39	1	7.55	0.0060
(2 vs 1) 40+	1	68.39	0.0000
(3 vs 1) 20-29	1	1.80	0.1798
(3 vs 1) 30-39	1	10.47	0.0012
(3 vs 1) 40+	1	159.89	0.0000
Joint	6	186.87	0.0000

	Contrast	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
group@agegroup						
(2 vs 1) 20-29	-.0058686	.0230533	-0.25	0.799	-.0510523	.039315
(2 vs 1) 30-39	-.0712194	.0259246	-2.75	0.006	-.1220308	-.0204081
(2 vs 1) 40+	-.1965602	.0237688	-8.27	0.000	-.2431461	-.1499742
(3 vs 1) 20-29	-.0284991	.0212476	-1.34	0.180	-.0701436	.0131454
(3 vs 1) 30-39	-.0861243	.0266137	-3.24	0.001	-.1382862	-.0339624
(3 vs 1) 40+	-.326247	.0258009	-12.64	0.000	-.3768159	-.2756781

The joint tests in the top table have been replaced by one-degree-of-freedom tests, one for each combination of the two reference comparisons and three levels of `agegroup`. The reference-category contrasts for `group` have been performed within levels of `agegroup`.

◁

Contrasts and the at() option

▷ Example 5

The `at()` option of `margins` is used to set predictors to particular values. When `at()` is used, contrasts are by default performed within each `at()` level:

```
. margins r.agegroup, at(group=(1/3))

Contrasts of adjusted predictions                      Number of obs = 3,000

Expression: Pr(outcome), predict()
1._at: group = 1
2._at: group = 2
3._at: group = 3

------------------------------------------------------------------
                         |         df        chi2     P>chi2
-------------------------+----------------------------------------
           agegroup@_at  |
    (30-39 vs 20-29) 1   |          1        6.94     0.0084
    (30-39 vs 20-29) 2   |          1        1.18     0.2783
    (30-39 vs 20-29) 3   |          1        3.10     0.0783
      (40+ vs 20-29) 1   |          1      173.42     0.0000
      (40+ vs 20-29) 2   |          1       57.77     0.0000
      (40+ vs 20-29) 3   |          1        5.12     0.0236
                Joint    |          6      266.84     0.0000
------------------------------------------------------------------

-----------------------------------------------------------------------
                     |            Delta-method
                     |   Contrast   std. err.     [95% conf. interval]
---------------------+-------------------------------------------------
        agegroup@_at |
(30-39 vs 20-29) 1   |   .0819713   .0311208      .0209757     .142967
(30-39 vs 20-29) 2   |   .0166206   .0153309     -.0134275    .0466686
(30-39 vs 20-29) 3   |   .0243462   .0138291     -.0027584    .0514507
  (40+ vs 20-29) 1   |   .3447797   .0261811      .2934658    .3960937
  (40+ vs 20-29) 2   |   .1540882   .0202722      .1143554     .193821
  (40+ vs 20-29) 3   |   .0470319   .0207774       .006309    .0877548
-----------------------------------------------------------------------
```

Our option `at(group=(1/3))` manipulates the values of `group` and is therefore not equivalent to `over(group)`. We see that the reference-category contrasts for `agegroup` have been performed within each `at()` level. For a similar example that uses the `._at` operator instead of the `at()` option, see *Contrasts of at() groups—discrete effects* in [R] **marginsplot**.

The default within behavior of `at()` may be changed to joint behavior with the `atjoint` suboption:

```
. margins r.agegroup, at(group=(1/3)) contrast(atjoint)

Contrasts of adjusted predictions                      Number of obs = 3,000
Model VCE: OIM

Expression: Pr(outcome), predict()
1._at: group = 1
2._at: group = 2
3._at: group = 3

------------------------------------------------------------------
                              |         df        chi2     P>chi2
------------------------------+-----------------------------------
                _at#agegroup  |
  (joint) (30-39 vs 20-29)    |          2        3.62     0.1641
    (joint) (40+ vs 20-29)    |          2       79.45     0.0000
                      Joint   |          4       83.88     0.0000
------------------------------------------------------------------
```

Now, the tests are performed jointly over the levels of `group`, the `at()` variable. The `atjoint` suboption is the analogue for `at()` of the `overjoint` suboption from example 3.

◁

▷ Example 6

What if we would like to apply a contrast operator, like `r.`, to the `at()` levels? It is not possible to specify the operator inside the `at()` option. Instead, we need a new suboption, `atcontrast()`:

```
. margins r.agegroup, at(group=(1/3)) contrast(atcontrast(r))

Contrasts of adjusted predictions                    Number of obs = 3,000
Model VCE: OIM

Expression: Pr(outcome), predict()
1._at: group = 1
2._at: group = 2
3._at: group = 3
```

	df	chi2	P>chi2
_at#agegroup			
(2 vs 1) (30-39 vs 20-29)	1	3.55	0.0596
(2 vs 1) (40+ vs 20-29)	1	33.17	0.0000
(3 vs 1) (30-39 vs 20-29)	1	2.86	0.0906
(3 vs 1) (40+ vs 20-29)	1	79.36	0.0000
Joint	4	83.88	0.0000

	Contrast	Delta-method std. err.	[95% conf.	interval]
_at#agegroup				
(2 vs 1) (30-39 vs 20-29)	-.0653508	.0346921	-.133346	.0026445
(2 vs 1) (40+ vs 20-29)	-.1906915	.0331121	-.25559	-.1257931
(3 vs 1) (30-39 vs 20-29)	-.0576252	.0340551	-.1243719	.0091216
(3 vs 1) (40+ vs 20-29)	-.2977479	.0334237	-.3632572	-.2322386

When we specify `contrast(atcontrast(r))`, `margins` will apply the `r.` reference-category operator to the levels of `group`, the variable specified inside `at()`. The default reference category is `1`, the lowest level of `group`.

Estimating treatment effects with margins

`margins` with the `contrast` option can also be used to estimate treatment effects in certain cases. A treatment effect represents the change in an outcome variable that is attributable to a particular event, controlling for all other factors that could affect the outcome. For example, we might want to know how a person's wage changes as a result of being in a union. Here the outcome variable is the person's wage, and the "event" is membership in a union. The treatment effect measures the difference in a person's wage as a result of being or not being in a union once we control for the person's educational background, level of experience, industry, and other factors.

In fact, Stata has an entire manual dedicated to estimators designed specifically for estimating treatment effects; see the *Stata Causal Inference and Treatment-Effects Estimation Reference Manual*. Here we show how `margins` can be used to estimate treatment effects using the regression-adjustment estimator when the conditional independence assumption is met; see [CAUSAL] **teffects intro**. Regression adjustment simply means that we are going to use a regression model to predict the outcome variable, controlling for treatment status and other characteristics. The conditional independence assumption implies that we have enough variables in our dataset so that once we control for them in our regression model, the outcomes one would obtain with and without treatment are independent of how treatment status is determined.

▷ Example 7: Regression adjustment with a binary treatment variable

nlsw88.dta contains women's wages (wage) in dollars per hour, a binary variable indicating their union status (union), years of experience (ttl_exp), and a variable, grade, indicating the number of years of schooling completed. We want to know how being in a union (the treatment) affects women's wages. Traditionally, a wage equation of the form

$$\ln \texttt{wage}_i = \beta_0 + \beta_1 \texttt{union}_i + \beta_2 \texttt{grade}_i + \beta_3 \texttt{ttl\_exp} + \beta_4 \texttt{ttl\_exp}^2 + \epsilon_i$$

would be fit. However, there are two shortcomings that we will improve upon. First, to avoid the problem of predicting the level of a log-transformed dependent variable, we will use poisson with the vce(robust) option to fit an exponential regression model; see Wooldridge (2010, sec. 18.2) for background on this approach. Second, the previous equation implies that factors other than union status have the same impact on wages for both union and nonunion workers. Regression-adjustment estimators allow all the variables to have different impacts depending on the level of the treatment variable, and we can accomplish that here using factor-variable notation. In Stata, we fit our model by typing

```
. use https://www.stata-press.com/data/r18/nlsw88
(NLSW, 1988 extract)

. poisson wage i.union##(c.grade c.ttl_exp##c.ttl_exp), vce(robust)
note: noncount dependent variable encountered; results correspond to an
      exponential-mean model rather than a poisson model.

Iteration 0:  Log pseudolikelihood = -4770.7957
Iteration 1:  Log pseudolikelihood = -4770.7693
Iteration 2:  Log pseudolikelihood = -4770.7693

Poisson regression                                     Number of obs =   1,876
                                                       Wald chi2(7)  = 1047.11
                                                       Prob > chi2   =  0.0000
Log pseudolikelihood = -4770.7693                      Pseudo R2     =  0.1195
```

wage	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
union						
Union	.8638376	.168233	5.13	0.000	.534107	1.193568
grade	.0895252	.0056874	15.74	0.000	.0783782	.1006722
ttl_exp	.0805737	.0114534	7.03	0.000	.0581255	.103022
c.ttl_exp# c.ttl_exp	-.0015502	.0004612	-3.36	0.001	-.0024541	-.0006463
union# c.grade						
Union	-.0310298	.0088259	-3.52	0.000	-.0483282	-.0137314
union# c.ttl_exp						
Union	-.0404226	.0230113	-1.76	0.079	-.085524	.0046788
union# c.ttl_exp# c.ttl_exp						
Union	.0011808	.0008428	1.40	0.161	-.0004711	.0028327
_cons	.017488	.0893602	0.20	0.845	-.1576547	.1926308

To see how union status affects wages, we can use `margins`:

```
. margins r.union, vce(unconditional)

Contrasts of predictive margins                        Number of obs = 1,876

Expression: Predicted number of events, predict()
```

	df	chi2	P>chi2
union	1	26.22	0.0000

	Contrast	Unconditional std. err.	[95% conf.	interval]
union (Union vs Nonunion)	1.004119	.1960944	.6197815	1.388457

The estimated contrast 1.004 indicates that on average, belonging to a union causes a woman's wage to be slightly more than a dollar higher than if she were not in the union. This estimated contrast is called the average treatment effect (ATE). Conceptually, we predicted the wage of each woman in the estimation sample assuming she was in a union and obtained the sample mean. We then predicted each woman's wage assuming she was not in a union and obtained that sample mean. The difference between these two sample means represents the ATE.

We obtain essentially the same results by using `teffects ra`:

```
. teffects ra (wage c.grade c.ttl_exp##c.ttl_exp, poisson) (union)

Iteration 0:   EE criterion =  2.611e-13
Iteration 1:   EE criterion =  1.112e-26

Treatment-effects estimation                   Number of obs     =      1,876
Estimator      : regression adjustment
Outcome model  : Poisson
Treatment model: none
```

wage	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
ATE						
union (Union vs Nonunion)	1.004119	.1960421	5.12	0.000	.619884	1.388355
POmean						
union Nonunion	7.346493	.1096182	67.02	0.000	7.131645	7.561341

The point estimates of the ATE are identical to those we obtained using `margins`, though the standard errors differ slightly from those reported by `margins`. The standard errors from the two estimators are, however, asymptotically equivalent, meaning they would coincide with a sufficiently large dataset. The last statistic in this output table indicates the untreated potential-outcome mean (untreated POM), which is the mean predicted wage assuming each woman did not belong to a union.

If we specify the `pomeans` option with `teffects ra`, we can obtain both the treated and the untreated POMs, which represent the predicted mean wages assuming all women were or were not in the union:

```
. teffects ra (wage c.grade c.ttl_exp##c.ttl_exp, poisson) (union), pomeans
Iteration 0:   EE criterion =  2.611e-13
Iteration 1:   EE criterion =  1.112e-26
Treatment-effects estimation                   Number of obs     =      1,876
Estimator      : regression adjustment
Outcome model  : Poisson
Treatment model: none

------------------------------------------------------------------------------
             |               Robust
        wage | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
POmeans      |
       union |
    Nonunion |   7.346493   .1096182    67.02   0.000     7.131645    7.561341
       Union |   8.350612   .1757346    47.52   0.000     8.006179    8.695046
------------------------------------------------------------------------------
```

Notice that the difference between these two POMs equals 1.004119, which is the ATE we obtained earlier.

◁

In some applications, the average treatment effect of the treated (ATET) is more germane than the ATE. For example, if the untreated subjects in the sample could not possibly receive treatment (perhaps because a medical condition precludes their taking an experimental drug), then considering the counterfactual outcome had those subjects taken the drug may not be relevant. In these cases, the ATET is a better statistic because it measures the effect of the treatment only for those subjects who actually did receive treatment. Like the ATE, the ATET involves computing predicted outcomes for each treatment level, obtaining the sample means, and computing the difference between those two means. Unlike the ATE, however, we use only observations corresponding to treated subjects.

▷ Example 8: Regression adjustment with a binary treatment variable (continued)

Here we calculate the ATET of union membership, first using `margins`. Because `teffects ra` overwrote our estimation results, we first quietly refit our `poisson` model. We then call `margins` to obtain the ATET:

```
. quietly poisson wage i.union##(c.grade c.ttl_exp##c.ttl_exp), vce(robust)
. margins r.union, subpop(union) vce(unconditional)
Contrasts of predictive margins                          Number of obs   = 1,876
                                                         Subpop. no. obs =   460
Expression: Predicted number of events, predict()

------------------------------------------------
             |         df        chi2     P>chi2
-------------+----------------------------------
       union |          1       18.86     0.0000
------------------------------------------------

--------------------------------------------------------------
                   |            Unconditional
                   |   Contrast   std. err.     [95% conf. interval]
-------------------+------------------------------------------
             union |
(Union vs Nonunion)|    .901419   .2075863      .4945574    1.308281
--------------------------------------------------------------
```

The key here was specifying the `subpop(union)` option to restrict `margin`'s computations to those women who are union members. The results indicate that being in the union causes the union members' wages to be about 90 cents higher than they would otherwise be.

To replicate these results using `teffects ra`, we include the `atet` option to obtain ATETs:

```
. teffects ra (wage c.grade c.ttl_exp##c.ttl_exp, poisson) (union), atet
Iteration 0:   EE criterion =  2.611e-13
Iteration 1:   EE criterion =  9.347e-27
Treatment-effects estimation                   Number of obs     =      1,876
Estimator      : regression adjustment
Outcome model  : Poisson
Treatment model: none
```

wage	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
ATET						
union (Union vs Nonunion)	.901419	.2075309	4.34	0.000	.4946658	1.308172
POmean						
union Nonunion	7.776417	.162121	47.97	0.000	7.458665	8.094168

We obtain the same point estimate of the effect of union status as with `margins`. As before, the standard errors differ slightly between the two estimators, but they are asymptotically equivalent. The output also indicates that among the women who are in a union, their average wage would be $7.78 if they were not in a union.

◁

❏ Technical note

One advantage of the ATET over the ATE is that the ATET can be consistently estimated with slightly weaker assumptions than are required to consistently estimate the ATE. See *Comparing the ATE and ATET* in *Remarks and examples* of [CAUSAL] **teffects intro advanced**.

❏

Both `margins` and `teffects` can estimate treatment effects using regression adjustment, so which should you use? In addition to regression adjustment, the `teffects` command implements other estimators of treatment effects; some of these estimators possess desirable robustness properties that we cannot replicate using `margins`. Moreover, all the `teffects` estimators use a common syntax and automatically present the estimated treatment effects, whereas we must first fit our own regression model and then call `margins` to obtain the treatment effects.

On the other hand, particularly with the `at()` option, `margins` gives us more flexibility in specifying our scenarios. The `teffects` commands allow us to measure the effect of a single binary or multinomial treatment, but we can have `margins` compute the effects of arbitrary interventions, as we illustrate in the next example.

▷ Example 9: Interventions involving multiple variables

Suppose we want to see how women's wages would be affected if we could increase each woman's education level by one year. That is, we want to measure the treatment effect of an additional year of schooling. We assume that if a woman attains another year of schooling, she cannot simultaneously work. Thus, an additional year of education implies her total work experience must decrease by a year. The flexible `at()` option of `margins` allows us to manipulate both variables at once:

```
. quietly poisson wage i.union##(c.grade c.ttl_exp##c.ttl_exp), vce(robust)
. margins, at((asobserved) _all)
> at(grade=generate(grade+1) ttl_exp=generate(ttl_exp-1))
> contrast(atcontrast(r._at))

Contrasts of predictive margins                    Number of obs = 1,876
Model VCE: Robust

Expression: Predicted number of events, predict()
1._at: (asobserved)
2._at: grade   =   grade+1
       ttl_exp = ttl_exp-1
```

	df	chi2	P>chi2
_at	1	58.53	0.0000

	Contrast	Delta-method std. err.	[95% conf.	interval]
_at (2 vs 1)	.3390392	.0443161	.2521813	.4258971

The first `at()` option instructs `margins` to obtain predicted wages for all women in the sample using their existing values for `grade` and `ttl_exp` and to record the mean of those predictions. The second `at()` option instructs `margins` to obtain the mean predicted wage under the counterfactual scenario where each woman's education level is increased by one year and total work experience is simultaneously decreased by one year. The `contrast()` option instructs `margins` to compute the difference between the two means. The output indicates that increasing education by one year, which will necessarily decrease work experience by the same amount, will cause the average wage to increase by about 34 cents per hour, a statistically significant amount.

◁

Conclusion

`margins, contrast` is a powerful command, and its abundance of suboptions may seem daunting. The suboptions are in the service of only three goals, however. There are three things that `margins, contrast` can do with a factor variable or a set of `at()` definitions:

1. Perform contrasts across the levels of the factor or set (as in example 1).
2. Perform a joint test across the levels of the factor or set (as in example 5).
3. Perform other tests and contrasts within each level of the factor or set (as in example 4).

The default behavior for variables specified inside `at()`, `over()`, and `within()` is to perform contrasts within groups; the default behavior for variables in the *marginlist* is to perform joint tests across groups.

◁

Stored results

`margins, contrast` stores the following additional results in `r()`:

Scalars

`r(k_terms)`	number of terms participating in contrasts

Macros

`r(cmd)`	`contrast`
`r(cmd2)`	`margins`
`r(overall)`	`overall` or empty

Matrices

`r(L)`	matrix of contrasts applied to the margins
`r(chi2)`	vector of χ^2 statistics
`r(p)`	vector of p-values corresponding to `r(chi2)`
`r(df)`	vector of degrees of freedom corresponding to `r(p)`

`margins, contrast` with the `post` option also stores the following additional results in `e()`:

Scalars

`e(k_terms)`	number of terms participating in contrasts

Macros

`e(cmd)`	`contrast`
`e(cmd2)`	`margins`
`e(overall)`	`overall` or empty

Matrices

`e(L)`	matrix of contrasts applied to the margins
`e(chi2)`	vector of χ^2 statistics
`e(p)`	vector of p-values corresponding to `e(chi2)`
`e(df)`	vector of degrees of freedom corresponding to `e(p)`

Methods and formulas

See *Methods and formulas* in [R] **margins** and *Methods and formulas* in [R] **contrast**.

Reference

Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: MIT Press.

Also see

[R] **contrast** — Contrasts and linear hypothesis tests after estimation

[R] **lincom** — Linear combinations of parameters

[R] **margins** — Marginal means, predictive margins, and marginal effects

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

[R] **margins, pwcompare** — Pairwise comparisons of margins

[R] **pwcompare** — Pairwise comparisons