New methods of interpretation using marginal effects for nonlinear models

Scott Long

Departments of Sociology and Statistics
Indiana University

EUSMEX 2016: Mexican Stata Users Group
Mayo 18, 2016

Road map for talk

Goals
1. Present new methods of interpretation using marginal effects
2. Show how to implement these methods with Stata

Outline
1. Statistical background
   - Binary logit model
   - Standard definitions of marginal effects
   - Generalizations to concept of marginal effects
2. Stata commands
   - Estimation
   - Post-estimation using margins and lincom
   - SPost13’s π* commands
3. Example modeling the occurrence of diabetes

Logit model

Outcome of probability or odds

\[ \pi(x) = \text{Prob}(y = 1 \mid x) \quad \text{and} \quad \Omega(x) = \pi(x) / [1 - \pi(x)] \]

Multiplicative in odds

\[ \Omega(x) = \frac{\pi(x)}{1 - \pi(x)} = \exp(x' \beta) \times \exp(\beta_0) \times \ldots \]

Odds ratio: multiplicative change in \( \Omega(x) \) for change in \( x \) holding other variables constant.

Nonlinear in probability

\[ \pi(x) = \frac{\exp(x' \beta)}{1 + \exp(x' \beta)} = \Lambda(x' \beta) = \Lambda(\beta_0 + \beta_1 x_1 + \ldots) \]

Discrete change: additive change in probability for change in \( x \) holding other variables constant.

Definition of discrete change

1. \( x_k \) changes from start to end
2. Remaining \( x \)'s held constant at specific values \( x = x^* \)
3. Discrete change \( DC(x_k) \)
   \[ \frac{\Delta \pi(x)}{\Delta x_k \text{ (start } \rightarrow \text{ end)}} = \pi(x_k = \text{end}, x = x^*) - \pi(x_k = \text{start}, x = x^*) \]
4. Interpretation
   For a change in \( x_k \) from start to end, the probability changes by \( DC(x_k) \), holding other variables at the specified values.
5. Everything that follows could be done using marginal changes
   \[ \frac{\partial \pi(x)}{\partial x_k} = \frac{\partial \Lambda(\beta_0 + \beta_1 x_1 + \ldots)}{\partial x_k} \]

Summarizing the effect of \( x_k \)

Since \( \Delta \pi / \Delta x_k \) depends on where it is evaluated, how can it be summarized?

Summary measures

DC at the mean: change at the center of the data

\[ DCM(x_k) = \frac{\Delta \pi(x = x^*)}{\Delta x_k \text{ (start } \rightarrow \text{ end)}} \]

For someone who is average on all variables, increasing \( x_k \) from start to end changes the probability by \( DCM(x_k) \).

Average DC: average change in estimation sample

\[ ADC(x_k) = \frac{1}{N} \sum_{i=1}^{N} \frac{\Delta \pi(x = x^*)}{\Delta x_k \text{ (start } \rightarrow \text{ end)}} \]

On average, increasing \( x_k \) from start to end changes the probability by \( ADC(x_k) \).

Generalized discrete change

My talk focuses on generalizing these standard measures
Variations in computing discrete change

Standard effects shown in black; generalized effects in red

Conditional and average change

- Conditional on specific values
- Averaged in the estimation sample
- Averaged in a subsample

Amount of change

- Constant change
- Proportional change
- Change as function of x’s
- Change of a component in a multiplicative measure

Number of variables changed

- One variable
- Two or more related variables

Stata requirements

1. Stata 14.1 (most things can be done with Stata 13)
2. search spost13_ado to install SPost13
3. search eusmex to download example, dataset, and slides

Stata commands

Steps in analysis using official Stata

1. Fit model using factor syntax
   
   ```plaintext
   logit depvar i.female c.age c.age#c.age
   ```

2. Store estimates using `estimates store Model`

3. Make predictions from regression using `margins, post`
   
   ```plaintext
   post
   ```
   
   replaces regression results with `margins` results

4. Estimate linear functions of predictions using `lincom`

5. `estimates restore Model` restores the regression estimates

Using SPost13

1. `mchange`, `mtable`, `mgen` and `mlincom` are SPost wrappers

2. They simplify things, but everything can be done without them

Modeling diabetes

Cross-sectional data from Health and Retirement Survey

```
use hrs-gme-analysis2, clear
(hr-gme-analysis2.dta | Health & Retirement Study CHE sample | 2016-04-08)
```

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Min</th>
<th>Max</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>diabetes</td>
<td>.205</td>
<td>0</td>
<td>1</td>
<td>Respondent has diabetes?</td>
</tr>
<tr>
<td>age</td>
<td>69.3</td>
<td>53</td>
<td>101</td>
<td>Age</td>
</tr>
<tr>
<td>bmi</td>
<td>27.9</td>
<td>10.6</td>
<td>82.7</td>
<td>Body mass index (weight/height^2)</td>
</tr>
<tr>
<td>weight</td>
<td>174.9</td>
<td>73</td>
<td>400</td>
<td>Weight in pounds</td>
</tr>
<tr>
<td>height</td>
<td>66.3</td>
<td>48</td>
<td>89</td>
<td>Height in inches</td>
</tr>
<tr>
<td>white</td>
<td>.772</td>
<td>0</td>
<td>1</td>
<td>Is white respondent?</td>
</tr>
<tr>
<td>female</td>
<td>.568</td>
<td>0</td>
<td>1</td>
<td>Is female?</td>
</tr>
<tr>
<td>hsdegree</td>
<td>.762</td>
<td>0</td>
<td>1</td>
<td>Has high school degree?</td>
</tr>
</tbody>
</table>

N=16,071

1Steve Heeringa generously provided the data used in Applied Survey Data Analysis (Heeringa et al., 2010). Complex sampling is not used in my analyses.

Two logit model specifications

1. Diabetes
   
   1.1 Given the diseases burden, small effects are substantively important
   
   1.2 With N=16,071 small effects are statistically significant

2. Two models that vary in how body mass is included

3. Model Mbmi uses the BMI index
   
   ```plaintext
   logit diabetes c.bmi ///
   i.white c.age#c.age i.female i.hsdegree
   ```

4. Model Mwt uses height and weight
   
   ```plaintext
   logit diabetes c.weight c.height ///
   i.white c.age#c.age i.female i.hsdegree
   ```

5. The estimates are...

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mbmi</th>
<th>Mwt</th>
</tr>
</thead>
<tbody>
<tr>
<td>bmi</td>
<td>1.1046*</td>
<td>1.1046*</td>
</tr>
<tr>
<td>weight</td>
<td>1.0165*</td>
<td>1.0165*</td>
</tr>
<tr>
<td>height</td>
<td>0.9299*</td>
<td>0.9299*</td>
</tr>
<tr>
<td>white</td>
<td>0.5412*</td>
<td>0.5313*</td>
</tr>
<tr>
<td>female</td>
<td>0.7191*</td>
<td>0.7191*</td>
</tr>
<tr>
<td>hsdegree</td>
<td>0.6711*</td>
<td>0.6711*</td>
</tr>
<tr>
<td>_cons</td>
<td>0.0000*</td>
<td>0.0000*</td>
</tr>
<tr>
<td>bic</td>
<td>14991.26</td>
<td>14982.03</td>
</tr>
</tbody>
</table>

Note: # significant at .05 level; * at the .001 level.
Average discrete change

1. After estimation I always run mchange

```
. estimates restore Mbmi
. mchange, amount(sd) // compute average discrete change
```

```
logit: Changes in Pr(y) | Number of obs = 16071

Change          p-value
bmi +SD          0.097  0.000
white vs Non-white -0.099  0.000
```

(output omitted)

2. Interpretation

*Increasing BMI by one standard deviation on average increases the probability of diabetes .097 (p < .001).*

*On average, the probability of diabetes is .099 less for white respondents than non-white respondents (p < .001).*

3. How were the DCs computed?

---

ADC for binary $x_k$: ADC(white)

1. ADC(white) is the difference in average probabilities

\[ ADC = \frac{1}{N} \sum_i \pi(white = 1, x = x_i) - \frac{1}{N} \sum_i \pi(white = 0, x = x_i) \]

2. Compute the two averages

```
. margins, at(white=0) at(white=1) post
```

| Expression            | Margin | Std. Err. | z    | P>|z|   | [95% Conf. Interval] |
|-----------------------|--------|-----------|------|-------|----------------------|
| _at                   |        |           |      |       |                      |
| 1._at : white = 0     | 0.2797806 | 0.0073107 | 38.27 | 0.000 | 0.265452 - 0.2941092 |
| 2._at : white = 1     | 0.1805306 | 0.0034215 | 52.76 | 0.000 | 0.1738245 - 0.1872367 |

3. Option post save the predictions to matrix e(b)

---

Tool: margins, at( ... ) and atmeans

1. By default, margins

1.1 Makes predictions for every case conditional on observed values

1.2 These conditional predictions are then averaged

2. Options allow counterfactual predictions

3. Average prediction imagining everyone is white

```
. margins, at(white=1)
```

4. Average predictions under two conditions

```
. margins, at(white=1) at(white=0)
```

5. Conditional prediction for someone white and average for other variables

```
. margins, at(white=1) atmeans
```

---

ADC for continuous $x_k$: ADC(bmi + sd)

4. The posted predictions from margins

```
. matlist e(b)
1.       2.
_y1       .2797806
1._at     0.1805306
```

5. lincom computes ADC(white) matching prior results

```
. lincom _b[2._at] - _b[1._at]
( 1) - 1bn._at + 2._at = 0
```

```
Coef. Std. Err. z    P>|z| [95% Conf. Interval]
(1)  -0.09925   0.0082362 -12.05 0.000 -0.1153927 -0.0831073
```

*On average, being white decreases the probability of diabetes by .099 (p < .001).*

---

Tool: margins, at( varnm = generate(exp) )

1. at( varnm = generate(exp) ) is powerful but poorly documented

2. Trivially, average prediction at observed values of bmi

```
. margins, at(bmi = gen(bmi))
```

3. Average prediction at the observed bmi plus 1

```
. margins, at(bmi = gen(bmi + 1))
```

4. Two average predictions

```
. margins, at(bmi = gen(bmi)) at(bmi = gen(bmi + 1))
```

5. Average at observed plus standard deviation

1) quietly sum bmi // summary statistics

```
2) local sd = r(sd) // retrieve standard deviation
```

```
3) margins, at(bmi = gen(bmi + `sd'))
```

1. Compute probabilities at observed bmi and observed + sd

```
. quietly sum bmi
. local sd = r(sd)
. margins, at(bmi = gen(bmi)) at(bmi = gen(bmi + `sd')) post
```

| Expression            | Margin | Std. Err. | z    | P>|z|   | [95% Conf. Interval] |
|-----------------------|--------|-----------|------|-------|----------------------|
| _at                   |        |           |      |       |                      |
| 1._at : bmi = bmi     | 0.2047165 | 0.0030338 | 67.48 | 0.000 | 0.1987704 - 0.2106627 |
| 2._at : bmi = bmi + 5.77032041238605 | 0.3017056 | 0.005199 | 58.03 | 0.000 | 0.2915159 - 0.3118954 |

6. ADC(bmi + sd)

```
. lincom _b[2._at] - _b[1._at]
( 1) - 1bn._at + 2._at = 0
```

```
Coef. Std. Err. z    P>|z| [95% Conf. Interval]
(1)  -0.09925   0.0082362 -12.05 0.000 -0.1153927 -0.0831073
```

*On average, increasing BMI by one standard deviation, about 6 points, increases the probability of diabetes by .097 (p < .001).*
**Tool: mlincom simplifies lincom**

1. `lincom` requires column names from `e(b)`
   ```
   . lincom _b[2._at] - _b[1._at]
   (1) = 1bn._at + 2._at = 0
   Coef. Std. Err. z P>|z| [95% Conf. Interval]
   (1)  .0969891  .0035648  27.21 0.000  .0900023  .1039759
   ```

2. `mlincom` uses column numbers in `e(b)` or rows in `margins` output
   ```
   . mlincom 2 - 1, stats(all)
   Expression: Pr(diabetes), predict()
   Pr(y)
   1 0.205
   2 0.271
   ```

3. Why use `mlincom`?
   ```
   . lincom (_b[2._at#1.white] - _b[1._at#1.white]) ///
   - (_b[2._at#0.white] - _b[1._at#0.white])
   mlincom (4-2) - (3-1)
   ```

**Generalized measures of discrete change**

1. DCM and ADC can be computed more easily with other commands
2. However, the commands showed are essential tools for computing generalized marginal effects
3. Examples of generalizations
   3.1 Proportional change in $x_k$
   3.2 Changing linked variables
   3.3 Distribution of effects
   3.4 Testing effects within a model
   3.5 Testing effects across models
   3.6 Testing effects across groups
   3.7 Changing a component of an interaction

**Tool: mtable wrapper for margins**

1. `margins` output is complete, not compact
2. `mtable` executes `margins` and simplifies the output and creates tables
   2.1 To list the `margins` commands used, add option `commands`
   2.2 To list `margins` and `mtable` output, add option `details`

**Proportional change in $x_k$: ADC(weight+25)**

1. Compute ADC(weight + 25)
   ```
   . estimates restore Mwt
   . mtable, at(weight = gen(weight)) at(weight = gen(weight + 25)) post
   Expression: Pr(diabetes), predict()
   Pr(y)
   1 0.205
   2 0.271
   ```

2. A simple change computes ADC(weight * 1.14)
   ```
   . estimates restore Mwt
   . mtable, at(weight = gen(weight)) at(weight = gen(weight * 1.14)) post
   Expression: Pr(diabetes), predict()
   Pr(y)
   1 0.205
   2 0.273
   ```

3. The effects are deceptively similar as shown below

**Proportional change in $x_k$: ADC(weight*1.14)**

2. A simple change computes ADC(weight * 1.14)
   ```
   . estimates restore Mwt
   . mtable, at(weight = gen(weight)) at(weight = gen(weight * 1.14)) post
   Expression: Pr(diabetes), predict()
   Pr(y)
   1 0.205
   2 0.273
   ```

2. To list `mtable` output, add option `details`

3. The effects are deceptively similar as shown below
Discrete change with linked variables

Mathematically linked variables
1. With polynomials multiple variables must change together
\[ \frac{\Delta \pi(x)}{\Delta \text{age}(50 - 60)} = \pi(\text{age}=60, \text{agesq}=60^2) - \pi(\text{age}=50, \text{agesq}=50^2) \]
2. With factor syntax margins handles this automatically

Substantively linked variables
1. Sometimes it makes sense to change multiple variables that are not mathematically linked
2. If two people have the same body mass, is the larger person more likely to have diabetes (the person who it taller and proportionally heavier)?
3. I compute an effect where height and weight change proportionally
4. Use \( \text{height} \) to predict \( \text{weight} \)
5. Use \( \text{at(\ldots=\text{gen()})} \) to change \( \text{height} \) and \( \text{weight} \) together

Linked variables: ADC(height, weight)

1. Regress \( \text{weight} \) on \( \text{height} \) and \( \text{height} \) squared
   \[ \text{regress weight c.height##c.height, noci} \]
   (output omitted)
   \[ \text{R-squared} = 0.2575 \]
   \[ \begin{array}{lcccc}
   & \text{weight} & \text{Coef.} & \text{Std. Err.} & \text{t} & \text{P>|t|} \\
   \text{height} & -6.338708 & 1.61073 & -3.94 & 0.000 \\
   \text{c.height#c.height} & .0855799 & .0120867 & 7.08 & 0.000 \\
   \_cons & 217.5991 & 53.5548 & 4.06 & 0.000 \\
   \end{array} \]

2. Save the estimates
   . scalar \( b0 = \_b[\_cons] \)
   . scalar \( b1 = \_b[\text{height}] \)
   . scalar \( b2 = \_b[\text{c.height#c.height}] \)
3. Weight can be predicted
   \[ \text{weight} = b0 + b1*\text{height} + b2*\text{height}^2 \]

Distribution of effects: limitations of summaries

1. ADC and DCM use averages
2. Average discrete change
   \[ \text{ADC}(x_1) = \frac{1}{N} \sum_i \left[ \frac{\Delta \pi}{\Delta (x_1| x = x_i)} \right] \]
3. Discrete change at the mean
   \[ \text{DCM}(x_1) = \frac{\Delta \pi}{\Delta (x_1| \bar{x} = \bar{x})} \]
4. Sometimes the averages distort the effect of a variables
5. Age has a large impact on diabetes, but ADC and DCM are small. Why?

Undocumented Tool: margins, generate()

1. margins, generate(stub) creates variables containing predictions for each observation
   . margins, generate(Prob)
   Predictive margins
   Number of obs = 16,071
   Expression: \( \text{Pr(diabetes)}, \text{predict()} \)
   Delta-method
   Margin Std. Err. z P>|z| [95% Conf. Interval]
   \_cons .2047166 .0030316 67.53 0.000 .1987747 .2106584
2. For details, help margins generate

Distribution of effects: ADC and DCM

Hypothetical data

![Distribution of effects: ADC and DCM](diagram)
1. To evaluate ADC(age) look at the distribution of DC(age)

2. Create a variable with the DC for each observation
   
   1a] margins, generate(PRage) ///
   1b] at(age = gen(age)) at(age = gen(age+10))

2a] gen DCage10 = PRage2 - PRage1

2b] lab var DCage10 "DC for 10 year increase in age"

3. Since age-squared was specified using factor syntax, when age is changed age#age is automatically changed

4. A histogram shows why ADC(age) is small

---

1. The average effect of age is small
2. The effect is large and negative for some people
3. The effect is large and positive for others

---

1. Early we computed ADC(weight+25) and ADC(weight*1.14)

2. The ADCs are similar but the distributions are quite different

---

1. Consider ADC(race) and ADC(bmi+sd)

2. Do the effects have the same size?

3. To answer this, the effects must be estimated simultaneously

---

1. ADC and DCM can be useful summaries, but in nonlinear models any summary measures can be misleading
2. The distribution of effects is valuable for assessing effects
3. This is simple with margins, generate()
4. Long and Freese (2014) show how do this in earlier versions of Stata

---

4. Merge the commands for ADC(white) and ADC(bmi)

4a] quietly sum bmi
   4b] local sd = r(sd)

4c] margins, at(white = 0) at(white = 1) ///
     at(bmi = gen(bmi)) at(bmi = gen(bmi + `sd´)) post

Predictive margins Number of obs = 16,071
Model VCE : OIM
Expression : Pr(diabetes), predict()
1._at : white = 0
2._at : white = 1
3._at : bmi = bmi
4._at : bmi = bmi + 5.770835041238605

Delta-method
Margin Std. Err. z P>|z| [95% Conf. Interval]
1._at .2797806 .0073107 38.27 0.000 .265452 .2941092
2._at .1805306 .0034215 52.76 0.000 .1738245 .1872367
3._at .2047166 .0030338 67.48 0.000 .1987704 .2106627
4._at .3017056 .005199 58.03 0.000 .2915159 .3118954
Comparing ADC(white) and ADC(bmi)

5. Compute ADCs and test equality
   . qui mlincom (2-1), rowname(ADC white)
   . qui mlincom (4-3), rowname(ADC bmi) add
   . mlincom (2-1) + (4-3), rowname(Sum of ADCs) add
   lincom pvalue ll ul

<table>
<thead>
<tr>
<th>ADC</th>
<th>pvalue</th>
<th>ll</th>
<th>ul</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADC white</td>
<td>-0.099</td>
<td>0.000</td>
<td>-0.115</td>
</tr>
<tr>
<td>ADC bmi</td>
<td>0.097</td>
<td>0.000</td>
<td>0.090</td>
</tr>
<tr>
<td>Sum of ADCs</td>
<td>-0.002</td>
<td>0.809</td>
<td>-0.021</td>
</tr>
</tbody>
</table>

6. Conclusion
   The health cost of being non-white is equivalent to a standard deviation increase in body mass ($p > .80$).

Comparing ADCs across models

1. Is ADC(female) the same across model specifications?

2. Tool: margins, dydx(female) computes DC(female) since i.female

3. Compute ADC(female) for two models separately
   . qui logit diabetes c.bmi i.female i.white c.age##c.age i.hdegree
   . qui mtable, dydx(female) rowname(ADC(female) with Mbmi) clear
   . qui logit diabetes c.weight c.height i.female i.white c.age##c.age i.hdegree
   . mtable, dydx(female) rowname(ADC(female) with Mwt) below

   Expression: Pr(diabetes), predict()

   . d Pr(y)
   | ADC(female) with Mbmi | -0.036 |
   | ADC(female) with Mwt  | -0.020 |

4. To test if effects are equal, they must be estimated simultaneously

   Tool: simultaneous model estimation with gsem

   1. gsem simultaneously fits multiple generalized linear models

   2. The obvious approach does not work since
      gsem ///
      (diabetes <- c.bmi i.female, logit) ///
      (diabetes <- c.weight c.height i.female, logit)

      is interpreted as
      gsem ///
      (diabetes <- c.bmi i.female c.weight c.height, logit)

   3. The solution is a cloned outcome for each model
      clonevar lbhbm = diabetes // outcome for Mbmi
      clonevar lhswt = diabetes // outcome for Mwt

Comparing ADC(female) across models

2. Simultaneously estimate ADC(female) for both models
   . margins, dydx(female) post

   Average marginal effects
   Number of obs = 16,071
   Model VCE : Robust
dy/dx v.r.t. : 1.female
1._predict : Predicted mean (Respondent has diabetes?), predict(pr
outcome(outcome(lbhbm)))
2._predict : Predicted mean (Respondent has diabetes?), predict(pr
outcome(lhswt))

   Delta-method
   dy/dx Std. Err. z  P>|z|  [95% Conf. Interval]
   1.female
      _predict 1 -.0360559 .0061773 -5.84 0.000 -.0481631 -.0239487
      _predict 2 -.0199213 .0089687 -2.22 0.026 -.0374997 -.0023429

   Note: dy/dx for factor levels is the discrete change from the base level.

   3. The estimates are identical to those estimate earlier

Comparing ADC(female) across models

4. Testing if the effects are equal
   . lincom 1-2, stats(all)

   | lincom se  zvalue  pvalue  ll    ul |
   |-------|--------|--------|-------|-------|
   | 1     | -0.016 | 0.006  | -2.526 | 0.012 | -0.029 | -0.004 |

5. Interpretation
   The effect of being female is significantly larger when body mass is measured with the BMI index then when height and weight are used to measure body mass ($p < .02$).
Comparing effects across models: summary

1. Jointly fitting models and estimating effects with margins is a general approach for comparing effects across models (Mize et al., 2009)
2. The `gsem` command
   - Fits GLM models only
   - margins is slow (grumble, grumble), but easy to use
3. Alternatively, the `suest` command
   - Fits a much wider class of models
   - margins is fast, but hard to use (grumble, grumble)
4. `suest` and `gsem` produce identical results
5. Specialized commands like `khb` (Kohler et al., 2011) are available

Comparing ADC across subsamples

1. An ADC is typically averaged over the entire sample
2. By averaging within groups, we can examine effects for different groups
   - Is the average effect of BMI the same for whites and non-whites?
3. To test if effects are equal across groups, we estimate the two effects simultaneously `margins, over()`

Tool: `margins, over()`

1. By default, margins averages all observations
2. Average for the non-white subsample
   ```stata
   margins if white==0, ///
   at(bmi = gen(bmi)) at(bmi = gen(bmi + `sd'))
   ```
3. Average for the white subsample
   ```stata
   margins if white==1, ///
   at(bmi = gen(bmi)) at(bmi = gen(bmi + `sd'))
   ```
4. Average for both subsamples simultaneously
   ```stata
   margins, over(white) ///
   at(bmi = gen(bmi)) at(bmi = gen(bmi + `sd'))
   ```

Comparing ADC(bmi) by race

1. To compute components for group specific ADC(bmi)
   ```stata
   . qui mlincom 4-2, clear rowname(White: ADC bmi)
   . mlincom 3-1, add rowname(Non-white: ADC bmi)
   ```
2. A second difference compares effects for the groups
   ```stata
   . mlincom (4-2) - (3-1), rowname(Difference: ADC bmi)
   ```
3. Interpretaion
   The average effect of BMI is significantly larger for non-whites than whites ($p < .001$).

Decomposing an effect

1. The BMI index measures relative weight
   $$\text{BMI} = \frac{703 \times \text{weight}^2}{\text{height}^2} = 703 \times \text{weight} \times \text{height}^{-1} \times \text{height}^{-1}$$
2. Why do this? DC(weight) is clearer to patients than DC(bmi)
Decomposing BMI: BMI is an interaction

1. Create components of BMI
   - `generate heightinv = 1/height`
   - `label var heightinv "1/height"`
   - `generate S = 703`
   - `label var S "scale factor to convert from metric"`

2. These models are identical
   - `logit diabetes c.S#c.weight#c.heightinv#c.heightinv ///`
   - `i.white c.age##c.age i.female i.hsdegree`
   - `estimates store MbmiFV`
   - `logit diabetes c.bmi i.white c.age##c.age i.female i.hsdegree`
   - `estimates store Mbmi`

3. The estimates are identical
   - | Variable | MbmiFV | Mbmi |
   - | c.S#c.weight# | 1.104553 | 1.104553 |
   - | c.heightinv# | 0.000 | 0.000 |
   - | c.heightinv | 1.104553 | 1.104553 |
   - | bmi | 0.000 | 0.000 |
   - | white | .5411742 | .5411742 |
   - | White | 0.000 | 0.000 |

Decomposing BMI: ADC(weight)

4. `margins` with factor syntax makes the rest easy
5. ADC(weight) in MbmiFV changes only weight
   - `. qui estimates restore MbmiFV`
   - `. mchange weight, amount(sd) delta(25)`
   - `logit: Changes in Pr(y) | Number of obs = 16071`  
   - `Expression: Pr(diabetes), predict(pr)`
   - `Change | p-value`
   - `weight +25 | 0.065 | 0.000`

Conclusions

Model interpretation and Stata

1. Too often interpretation ends with estimated coefficients
   - Interpretation using predictions is more informative
   - I think of regression coefficients as “nuisance parameters”
2. Methods of interpretation must be practical
   - `margins` makes hard things easy, very hard things merely hard

Which method of interpretation?

1. `mchange` makes it easy make marginal effects a routine part of analysis; marginal effects are almost always more useful than odds ratios
2. Generalized marginal effects can be tailored to your research
3. But, marginal effects might not be the best method of interpretation
4. Tables and plots might be more useful (Long and Freese, 2014) and are easy with `margins` and the `m*` commands
5. The best interpretation is motivated by your substantive question

Thanks to many people

Thank you for listening

Collaborators
Parts of this work were developed with Long Doan, Jeremy Freese, Trent Mize, and Sarah Mustillo. Jeff Pitblado and David Drukker provided valuable help. Mistakes are my own.

Relevant publications
There is a large literature on marginal effects and interpreting models. Long and Freese (2014) include many citations. The references directly related to this presentation are given below.

Bibliography

Cameron, A. C., and P. K. Trivedi. 2010. Microeconometrics using Stata. Revised ed. College Station, Tex.: Stata Press.
**Additional examples**

1. Comparing ADC(weight) across models
2. Discrete change with polynomials
3. Comparing ADCs across models with `suest`
4. Comparing groups: outcomes and marginal effects
5. Computing DCMs
6. Comparing DCRs

---

**Comparing ADC(weight) across models**

4. To compare ADC(weight) requires joint estimation

```
. clonevar lhsbmi = diabetes
. clonevar lhswt = diabetes
. gsem
> (lhsbmi <- c.S#c.weight#c.heightinv#c.heightinv ///
> 1.white c.age##c.age i.female i.hsdegree, logit) ///
> (lhswt <- c.weight c.height ///
> 1.white c.age##c.age i.female i.hsdegree, logit) ///
> , vce(robust)
```

Generalized structural equation model
Number of obs = 16,071
Response : lhsbmi
Family : Bernoulli
Link : logit
Response : lhswt
Family : Bernoulli
Link : logit
Log pseudolikelihood = -14914.007
(output omitted)

---

**Comparing ADC(weight) across models**

5. Computing the average predictions for both equations

```
. margins, at(weight=gen(weight)) at(weight=gen(weight+25)) post
```

Predictive margins
Number of obs = 16,071
Model VCE : Robust
```
1._predict : Predicted mean (Diabetes?), predict(pr outcome(lhsbmi))
2._predict : Predicted mean (Diabetes?), predict(pr outcome(lhswt))
1._at : weight = weight
2._at : weight = weight+25
```

| Delta-method | Margin | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|--------------|--------|-----------|-------|------|---------------------|
| _predict#at  | 1 1    | 0.2047166 | 0.0030419 | 67.30 | 0.000 | 0.1987546 - 0.2106786 |
|              | 1 2    | 0.2701404 | 0.0044591 | 60.58 | 0.000 | 0.2614007 - 0.27888 |
|              | 2 1    | 0.2047166 | 0.0030394 | 67.35 | 0.000 | 0.1987595 - 0.2106737 |
|              | 2 2    | 0.271305  | 0.0044054 | 61.58 | 0.000 | 0.2626705 - 0.2799394 |
```

---

**Comparing ADC(weight) in two models**

6. ADC(weight) for each model and their difference

```
. qui mlincom 2-1, rowname(Mbmi ADC) clear
. qui mlincom 4-3, rowname(Mwt ADC) add
. mlincom (4-3) - (2-1), rowname(Difference) add
```

<table>
<thead>
<tr>
<th>lincom</th>
<th>prob</th>
<th>ll</th>
<th>ul</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mbmi ADC</td>
<td>0.065</td>
<td>0.000</td>
<td>0.061</td>
</tr>
<tr>
<td>Mwt ADC</td>
<td>0.067</td>
<td>0.000</td>
<td>0.062</td>
</tr>
<tr>
<td>Difference</td>
<td>0.001</td>
<td>0.029</td>
<td>0.000</td>
</tr>
</tbody>
</table>

7. Conclusion

*The effect of weight on diabetes are nearly identical whether body mass is measured with BMI or with height and weight (p = .03).*

---

**Discrete change with polynomials**

1. With polynomials multiple variables **must** change together
2. For example,

\[
\frac{\Delta \pi(x)}{\Delta age(50 \rightarrow 60)} = \pi(age=60, \text{agesq}=60^2) - \pi(age=50, \text{agesq}=50^2)
\]
3. This can be computed two ways
   3.1 Automatically with factor syntax
   3.2 Explicitly with `at(...) = gen(...)`)
Discrete change with polynomials

1. With \( x \) and \( x^2 \) only values on the blue curve are mathematically possible

2. Changes in the probability reflect linked changes in \( x \) and \( x^2 \)

---

**Tool:** factor notation for polynomials

**Without factor notation**

1. Create the squared term
   
   ```
   generate agesq = age * age
   ```

2. Then fit
   
   ```
   logit diabetes c.age c.agesq ...
   ```

**With factor notation**

1. Fit the model
   
   ```
   logit diabetes c.age##c.age ...
   ```

2. \( c.age##c.age \) automatically
   1. Adds \( c.age \) to the model
   2. Creates \( c.age\#c.age \equiv age*age \equiv agesq \)
   3. Adds \( c.age\#c.age \) to the model

3. When \( c.age \) changes, margins automatically changes \( c.age\#c.age \)

---

Discrete change with age & \( age^2 \)

Correct ADC(age) with factor & \( age^2 \)

1. \( age \) and \( age\#age \) automatically change together

   ```
   logit diabetes c.age##c.age c.bmi i.white i.female i.hsdegree, or
   ```

   ```
   . mtable, dydx(female) rowname(ADC(female) with Mbmi) clear
   ```

   ```
   logit diabetes c.weight c.height i.female i.white c.age##c.age i.hsdegree
   ```

   ```
   . mtable, dydx(female) rowname(ADC(female) with Mwt) below
   ```

3. To test if they are equal, the effects must be estimated simultaneously

---

Comparing ADCs across models with suest

1. Does the effect of a variable change with model specification?

2. Computing ADC(female) for two models

   ```
   qui logit diabetes c.bmi i.white i.female c.age##c.age i.hsdegree
   ```

   ```
   . estimate store Mbmi
   ```

   ```
   qui logit diabetes c.weight c.height i.female i.white c.age##c.age i.hsdegree
   ```

   ```
   . estimate store Mwt
   ```

   ```
   . mtable, dydx(female) rowname(ADC(female) with Mbmi) clear
   ```

   ```
   . mtable, dydx(female) rowname(ADC(female) with Mwt) below
   ```

3. Why use at(gen()) instead of factor syntax

   1. at(gen()) does many things that factor syntax cannot do (gripe)
Comparing effects across models: ADC(female)  
Joint estimation with `suest`

4. The stored estimates are combined and stored
   
   . `suest Mbmi Mwt, noci`
   
   Simultaneous results for `Mbmi`, `Mwt`

   Number of obs = 16,071

<table>
<thead>
<tr>
<th></th>
<th>Robust</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><code>Mbmi_diabetes</code></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>bmi</code></td>
<td>.099441</td>
<td>.003747</td>
<td>.000</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td><code>white</code></td>
<td>-.634014</td>
<td>.048092</td>
<td>-9.20</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td><code>Mwt</code></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>weight</code></td>
<td>.0163568</td>
<td>.0005901</td>
<td>27.72</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td><code>height</code></td>
<td>-.0726272</td>
<td>.0078904</td>
<td>-9.20</td>
<td>.000</td>
<td></td>
</tr>
</tbody>
</table>

. `qui estimates store _Msuest`

Second differences with `suest`

Comparing ADCs across models: ADC(female)

6. With `suest`, `margins` computes \( \hat{x}'\hat{\beta} \), but we need \( \hat{\pi}(x) = \Lambda(\hat{x}'\hat{\beta}) \)

7. Option `predict(equation(Mbmi_diabetes))` computes \( \hat{x}'\hat{\beta} \) for `Mbmi`

8. The logistic CDF function `logistic()` transforms \( \hat{x}'\hat{\beta} \) to \( \hat{\pi}(x) \)

9. The expression for \( \hat{\pi}(x) \) is
   
   `expression(logistic(predict(equation(Mbmi_diabetes)))`

10. To make code easier, save expressions for `Mbmi` and `Mwt` in locals
    
    `local _EXPR_Mbmi equation(diabetes Mbmi)`
    
    `local _EXPR_Mwt equation(diabetes Mwt)`

11. The rest is “easy”

## Tool: equation, predict, and expression

### ADC with `suest`

1. The two stored models are `equation` in the `suest` model
   
   `Mbmi` becomes `equation(diabetes_Mbmi)`
   
   `Mwt` becomes `equation(diabetes_Mwt)`

2. With logit, `margins` by default computes the “expression” for predicted probabilities
   
   Expression: \( Pr(diabetes), predict() \)

3. With `suest`, `margins` only computes \( x'\beta \)
   
   Expression: Linear prediction, `predict()`

4. Sadly, `margins`, `predict(pr)` does not work with `suest`

5. The solution is the `expression()` option

### Comparing ADCs across models: ADC(female)

2. For model `Mbmi`
   
   `. mtable, expression(`_EXPR_Mbmi´) at(female=1) at(female=0) post`  
   
   Expression: `logistic(predict(equation(Mbmi_diabetes)))`

3. The estimates match those from the individual models; standard errors are robust

## Comparing ADCs across models: ADC(female)

Second differences with `suest`

1. The ADCs from the two models are
   
   \[
   \text{ADC}_{\text{Mbmi}} = \pi_{\text{Mbmi}}(\text{female} = 1, x) - \pi_{\text{Mbmi}}(\text{female} = 0, x)
   \]
   
   \[
   \text{ADC}_{\text{Mwt}} = \pi_{\text{Mwt}}(\text{female} = 1, x) - \pi_{\text{Mwt}}(\text{female} = 0, x)
   \]

2. Since `margins` can’t compute these in one step, we compute the parts
   
   \[
   \hat{\pi}_{\text{Mbmi}}(\text{female} = 0, x) - \hat{\pi}_{\text{Mbmi}}(\text{female} = 0, x)
   \]
   
   \[
   \hat{\pi}_{\text{Mbmi}}(\text{female} = 1, x) - \hat{\pi}_{\text{Mbmi}}(\text{female} = 1, x)
   \]

3. Subtracting these is the second difference we want to test
   
   \[
   \text{ADC}_{\text{Mbmi}} - \text{ADC}_{\text{Mwt}} = \hat{\pi}_{\text{Mbmi}}(\text{female} = 1, x) - \hat{\pi}_{\text{Mbmi}}(\text{female} = 0, x) - [\hat{\pi}_{\text{Mbmi}}(\text{female} = 1, x) - \hat{\pi}_{\text{Mbmi}}(\text{female} = 0, x)]
   \]

4. The results from `margins` follow

5. Using the locals defined earlier
   
   `. mtable, expression(`_EXPR_Mbmi´-`_EXPR_Mwt´)`
   
   Expression: `logistic(predict(equation(Mbmi_diabetes)))`

6. The 2nd difference is
   
   `. mlincom` `1-2`, title(`Ho: ADC female equal for Mwt & Mbmi`)` 
   
   `Ho: ADC female equal for m_wt & Mbmi`
   
   `. mlincom` `1-2`, title(`Ho: ADC female equal for Mwt & Mbmi`)` 
   
   `Ho: ADC female equal for m_wt & Mbmi`  
   
   `. mtable, expression(`_EXPR_Mwt´) at(female=1) at(female=0) post`  
   
   `. qui mtable, expression(`_EXPR_Mw~t´) at(female=1) at(female=0) post`  
   
   `. qui mtable, expression(`_EXPR_Mw~t´) at(female=1) at(female=0) post`  
   
   `. qui mlincom` `1-2`, title(`Ho: ADC female equal for Mwt & Mbmi`)` 
   
   `. qui mlincom` `1-2`, title(`Ho: ADC female equal for Mwt & Mbmi`)`

7. Interpretation
   
   The effect of being female is significantly larger when body mass is measured with BMI than with weight and height (\( p < .02 \)).
Comparing effects across models: summary

1. Jointly fitting models and computing effects with margins is a general approach for comparing effects across models (Mize et al., 2009)

2. \texttt{gsem}
   2.1 Fits generalized linear models only
   2.2 \texttt{margins} is slow (grumble, grumble), but easy to use

3. \texttt{suest}
   3.1 Fits a much wider class of models
   3.2 \texttt{margins} is fast, but hard to use (grumble, grumble)

4. \texttt{suest} and \texttt{gsem} produce identical results

Comparing groups: outcomes and effects

Group differences can be examined two ways

1. Differences in probabilities
   \[ H_0: \pi_W(x) = \pi_N(x) \]
   Is the probability of diabetes the same for white and non-white respondents who have the same characteristics?

2. Differences in marginal effects
   \[ H_0: \frac{\Delta \pi_W}{\Delta x_k} = \frac{\Delta \pi_N}{\Delta x_k} \]
   Is the effect of \( x_k \) the same for whites and non-whites?

3. These dimensions of difference are shown in the next graph

Comparing groups: model estimation

1. Factor syntax allows coefficients to differ by \texttt{white}
   \[
   \text{logit diabetes \ i.white ///}
   \text{\ i.white#(i.female i.hsdegree c.age##c.age c.bmi), nocon}
   \]

2. This is equivalent to simultaneously estimating
   \[
   \text{logit diabetes \ i.female i.hsdegree c.age##c.age c.bmi if white==1}
   \text{logit diabetes \ i.female i.hsdegree c.age##c.age c.bmi if white==0}
   \]

3. Resulting in these estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>Whites</th>
<th>Nonwhites</th>
</tr>
</thead>
<tbody>
<tr>
<td>female</td>
<td>0.713</td>
<td>1.024</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.755</td>
</tr>
<tr>
<td>hsdegree</td>
<td>0.706</td>
<td>0.743</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>age</td>
<td>1.278</td>
<td>1.369</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Comparing groups: outcomes and effects

Hypothetical data

Group differences in probabilities by age

1. Compute DC(white) at different ages
   \[
   \text{.stable, dydx(white) at(age=(55(10)85)) atmeans stats(est p)}
   \]
   Expression: Pr(diabetes), predict()

| age | d Pr(y) | p    | p<DC(white | age=55) | p<DC(white | age=65) | p<DC(white | age=75) | p<DC(white | age=85) |
|-----|---------|------|------------|----------|------------|----------|------------|----------|------------|
| 1   | 55      | -0.078| 0.000      | <=   | <=   | <=   |
| 2   | 65      | -0.124| 0.000      | <=   | <=   | <=   |
| 3   | 75      | -0.129| 0.000      | <=   | <=   | <=   |
| 4   | 85      | -0.092| 0.000      | <=   | <=   | <=   |

Specified values of covariates

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>white</td>
<td>white</td>
<td>female</td>
<td>hsdegree</td>
<td>bmi</td>
</tr>
<tr>
<td>0.228</td>
<td>0.772</td>
<td>0.568</td>
<td>0.762</td>
<td>27.9</td>
</tr>
</tbody>
</table>

2. Example of interpretation
   For average respondents who are 55, the probability of diabetes is significantly larger for non-whites than whites (\( p<.01 \)).

3. Graphically we can show effects at multiple ages
Group differences in probabilities by age

A: Probabilities

B: DCR(race)

Note: these plots can be computed with mgen or marginsplot

Group differences in effects: summary

Comparing ADCs
1. Group differences in ADCs are determined by two things
   1.1 Group differences in the probability curves
   1.2 Group differences in distribution of variables

Comparing DCRs
1. Group differences in DCRs are determined by two things
   1.1 Group differences in the probability curves
   1.2 The specific location where they are evaluated
2. They do not depend on group differences in the distribution of variables

Which to use?
1. The answer depends on what you want to know?

Group differences in ADC(bmi + 5)

1. To compute ADC(bmi + 5) by race
   . stable, over(white) at(bmi = gen(bmi)) at(bmi = gen(bmi+5)) post
   Expression: Pr(diabetes), predict()

<table>
<thead>
<tr>
<th>Pr(y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.white#c.1</td>
</tr>
<tr>
<td>0.310</td>
</tr>
<tr>
<td>1.white#c.1</td>
</tr>
<tr>
<td>0.174</td>
</tr>
<tr>
<td>0.white#c.2</td>
</tr>
<tr>
<td>0.391</td>
</tr>
<tr>
<td>1.white#c.2</td>
</tr>
<tr>
<td>0.257</td>
</tr>
<tr>
<td>. qui mlcom 3-1, rowname(ADC(bmi) non) stats(est p) clear</td>
</tr>
<tr>
<td>. qui mlcom 4-2, rowname(ADC(bmi) wht) stats(est p) add</td>
</tr>
<tr>
<td>. mlcom (4-2) - (3-1), rowname(Diff) stats(est p) add</td>
</tr>
</tbody>
</table>

   lincom  pvalue
   ADC(bmi) non  0.082 0.000
   ADC(bmi) wht  0.083 0.000
   Difference    0.002 0.826

2. Conclusion
   The average effects of BMI are not significantly different for whites and non-whites (p=.83).

Group differences in DCR(age + 10)

1. ADC(age) might not be useful due to nonlinearity
2. We compare DCR(age+10) at different ages
   2.1 Other variables are held at sample means
   2.2 Group specific means could be used (Long and Freese, 2014)
3. For example, DCR(age + 10) at 55
   stable, atmeans post /// at(age=55) at(age=65 white=0) /// at(age=65) at(age=65 white=1)(
   mlcom 3-1, rowname(DC nonwhite) stats(est p) clear
   mlcom 4-2, rowname(DC white) stats(est p) add
   mlcom (4-2) - (3-1), rowname(Difference) stats(est p) add

4. And so on, with the following results
5. DCRs show group differences in effect of age at different ages

<table>
<thead>
<tr>
<th>lincom</th>
<th>pvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>55: DC non</td>
<td></td>
</tr>
<tr>
<td>0.110 0.000</td>
<td></td>
</tr>
<tr>
<td>DC white</td>
<td></td>
</tr>
<tr>
<td>0.064 0.000</td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td></td>
</tr>
<tr>
<td>-0.046 0.001</td>
<td></td>
</tr>
<tr>
<td>70: DC non</td>
<td></td>
</tr>
<tr>
<td>0.001 0.940</td>
<td></td>
</tr>
<tr>
<td>DC white</td>
<td></td>
</tr>
<tr>
<td>0.018 0.001</td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td></td>
</tr>
<tr>
<td>0.017 0.180</td>
<td></td>
</tr>
<tr>
<td>85: DC non</td>
<td></td>
</tr>
<tr>
<td>-0.109 0.000</td>
<td></td>
</tr>
<tr>
<td>DC white</td>
<td></td>
</tr>
<tr>
<td>-0.049 0.000</td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td></td>
</tr>
<tr>
<td>0.060 0.003</td>
<td></td>
</tr>
</tbody>
</table>

6. The differences in DCRs do not depend on group differences in the distribution of age or other variables
DCM for continuous $x_k$: DCM(bmi + sd)

Discrete change at the mean

1. Let bmi increase from mean(bmi) to mean(bmi) + sd(bmi)
   - qui sum bmi
   - local mn = r(mean)
   - local mnplus = r(mean) + r(sd)

2. Option atmeans holds other variables at their means
   - margins, atmeans at(bmi = `mn´) at(bmi = `mnplus´) post

Expression : Pr(diabetes), predict()

1._at : bmi = 27.9
   0.white = .228 (mean)
   1.white = .772 (mean)
   age = 69 (mean)
   0.female = .432 (mean)
   1.female = .568 (mean)
   0.hsdegree = .238 (mean)
   1.hsdegree = .762 (mean)

<continued>

DCM for continuous $x_k$: DCM(bmi + sd)

1. Let bmi increase from mean(bmi) to mean(bmi) + sd(bmi)
   - qui sum bmi
   - local mn = r(mean)
   - local mnplus = r(mean) + r(sd)

2. Option atmeans holds other variables at their means
   - margins, atmeans at(bmi = `mn´) at(bmi = `mnplus´) post

Expression : Pr(diabetes), predict()

1._at : bmi = 33.7
   0.white = .228 (mean)
   1.white = .772 (mean)
   age = 69 (mean)
   0.female = .432 (mean)
   1.female = .568 (mean)
   0.hsdegree = .238 (mean)
   1.hsdegree = .762 (mean)

Delta-method
Margin Std. Err. z P>|z| [95% Conf. Interval]
1._at 1 .2097641 .0045531 46.07 0.000 .2008401 .2186881
2 .3202789 .0066246 48.35 0.000 .307295 .3332628

3. For complex models the output gets very long, so mtable was written.

Tool: mtable wrapper for margins

1. margins output is complete, not compact
2. mtable executes margins and simplifies the output (and more)
   2.1 To see the margins commands being used, add option commands
   2.2 To see margins and mtable output, add option details

DCM for continuous $x_k$: DCM(bmi + sd)

1. Let bmi increase from mean(bmi) to mean(bmi) + sd(bmi)
   - qui sum bmi
   - local mn = r(mean)
   - local mnplus = r(mean) + r(sd)

2. Option atmeans holds other variables at their means
   - margins, atmeans at(bmi = `mn´) at(bmi = `mnplus´) post

Expression : Pr(diabetes), predict()

1._at : bmi = 27.9
   0.white = .228 (mean)
   1.white = .772 (mean)
   age = 69 (mean)
   0.female = .432 (mean)
   1.female = .568 (mean)
   0.hsdegree = .238 (mean)
   1.hsdegree = .762 (mean)

DCM for continuous $x_k$: DCM(bmi + sd)

1. Let bmi increase from mean(bmi) to mean(bmi) + sd(bmi)
   - qui sum bmi
   - local mn = r(mean)
   - local mnplus = r(mean) + r(sd)

2. Option atmeans holds other variables at their means
   - margins, atmeans at(bmi = `mn´) at(bmi = `mnplus´) post

Expression : Pr(diabetes), predict()

1._at : bmi = 33.7
   0.white = .228 (mean)
   1.white = .772 (mean)
   age = 69 (mean)
   0.female = .432 (mean)
   1.female = .568 (mean)
   0.hsdegree = .238 (mean)
   1.hsdegree = .762 (mean)

Delta-method
Margin Std. Err. z P>|z| [95% Conf. Interval]
1._at 1 .2097641 .0045531 46.07 0.000 .2008401 .2186881
2 .3202789 .0066246 48.35 0.000 .307295 .3332628

3. For complex models the output gets very long, so mtable was written.

Tool: mtable wrapper for margins

1. margins output is complete, not compact
2. mtable executes margins and simplifies the output (and more)
   2.1 To see the margins commands being used, add option commands
   2.2 To see margins and mtable output, add option details

Comparing DCRs

1. Is the effect of age significantly different at different ages?

Comparing DCR(age) at different ages

1. Is the effect of age significantly different at different ages?

2. Compute probabilities at four ages with other variables at means
   - mtable, at(age=(60(10)90)) post atmeans

Expression: Pr(diabetes), predict()

1. age Pr(y)
   1 60 .150
   2 70 .213
   3 80 .227
   4 90 .183

Specified values of covariates
   1. bmi white female hsdegree
      current .772 69.3 .568 ..76

For someone who is average, increasing BMI by one standard deviation increases the probability of diabetes by .111 (p < .001).
Comparing DCR(age) at different ages

4. Test differences in DCRs
   . mlincom (2-1) - (3-2), add rowname(DCR60 - DCR70)
   . mlincom (2-1) - (4-3), add rowname(DCR60 - DCR80)
   . mlincom (3-2) - (4-3), add rowname(DCR70 - DCR80)

5. Summarizing
   . mlincom, twidth(14)

<table>
<thead>
<tr>
<th></th>
<th>lincom</th>
<th>pvalue</th>
<th>ll</th>
<th>ul</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCR60</td>
<td>0.063</td>
<td>0.000</td>
<td>0.054</td>
<td>0.073</td>
</tr>
<tr>
<td>DCR70</td>
<td>0.014</td>
<td>0.004</td>
<td>0.004</td>
<td>0.023</td>
</tr>
<tr>
<td>DCR80</td>
<td>-0.043</td>
<td>0.000</td>
<td>-0.061</td>
<td>-0.026</td>
</tr>
<tr>
<td>DCR60 - DCR70</td>
<td>0.049</td>
<td>0.000</td>
<td>0.037</td>
<td>0.062</td>
</tr>
<tr>
<td>DCR60 - DCR80</td>
<td>0.107</td>
<td>0.000</td>
<td>0.083</td>
<td>0.130</td>
</tr>
<tr>
<td>DCR70 - DCR80</td>
<td>0.057</td>
<td>0.000</td>
<td>0.046</td>
<td>0.069</td>
</tr>
</tbody>
</table>

6. Interpretation
   The effects of a ten-year increase in age are significantly different at ages 60, 70, and 80 ($p < .001$).

The end

No more examples!