What does your model say? It may depend on who is asking

David M. Drukker

Executive Director of Econometrics
Stata

Stata Conference
Chicago
28 July 2016

Outline

I define and contrast conditional-on-covariate inference with population-averaged inference

College success data

- Simulated data on a college-success index (`csuccess`) on 1,000 students that entered an imaginary university in the same year
- `iexam` records each student's grade on the final from a mandatory short course that taught study techniques and new material attending prior to staring
- `sat` is combined math and verbal SAT score, recorded in hundreds of points
- `hgpa` is high-school grade-point average
- Want effect of the `iexam` score
- Include an "interaction term" `it=iexam/(hgpa^2)`
    - allows for the possibility that `iexam` has a smaller effect for students with a higher `hgpa`

```
. regress csuccess hgpa sat iexam it, vce(robust)
Linear regression                               Number of obs     =      1,000
                                                F(4, 995)         =     384.34
                                                Prob > F          =     0.0000
                                                R-squared         =     0.5843
                                                Root MSE          =     1.3737
```

csuccess	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
hgpa	.7030099	.178294	3.94	0.000	.3531344	1.052885
sat	1.011056	.0514416	19.65	0.000	.9101095	1.112002
iexam	.1779532	.0715848	2.49	0.013	.0374788	.3184276
it	5.450188	.3731664	14.61	0.000	4.717904	6.182471
_cons	-1.434994	1.059799	-1.35	0.176	-3.514692	.644704

The estimated conditional mean function

$$\widehat{\mathbf{E}}[\texttt{csuccess}|\texttt{hgpa}, \texttt{sat}, \texttt{iexam}]$$
$$= .70\texttt{hgpa} + 1.01\texttt{sat} + 0.18\texttt{iexam} + 5.45\texttt{iexam}/(\texttt{hgpa}^2) - 1.43$$

produces estimates of the mean of csuccess for given values of hgpa, sat, iexam

- My model of csuccess for given values of hgpa, sat, iexam is

$$\mathbf{E}[\texttt{csuccess}|\texttt{hgpa},\texttt{sat},\texttt{iexam}]$$
$$= \beta_1\texttt{hgpa} + \beta_2\texttt{sat} + \beta_3\texttt{iexam} + \beta_4\texttt{iexam}/(\texttt{hgpa}^2) + \beta_0$$

- Differences in $\mathbf{E}[\texttt{csuccess}|\texttt{hgpa},\texttt{sat},\texttt{iexam}]$ resulting from an everything-else-held-constant change of hgpa, sat, or iexam define causal effects
- This effect exists without reference to how the parameters are estimated
    - You tell me the values of the covariates specifying the everything-else-held-constant change and I can compute the effect
- Pluging in any consistent estimates of β_0, β_1, β_2, β_3, and β_4, produces consistent estimates of the effects
  How these estimates were computed has no bearing on the definition or the interpretation of the effects

Skip: Only discuss if questions require

- The derivation of regression adjustment in the modern causal inference literature uses this effect definition
    - This literature does not challenge that everything-else-held-constant changes in a well-specified conditional mean function define effects
      Rather
      it is about what are the exogeity assumptions and functional form assumptions that produce a well-specified conditional mean function
    - See Imbens (2004), Cameron and Trivedi (2005, chapter 2.7), Imbens and Wooldridge (2009), and Wooldridge (2010, chapters 2 and 21)

Effect of a 100-point increase in SAT

Because sat is measured in hundreds of points, the effect of a 100-point increase in sat is estimated to be

$$\begin{aligned}
&\widehat{\mathbf{E}}[\texttt{csuccess}|\texttt{hgpa}, (\texttt{sat}+1), \texttt{iexam}] - \widehat{\mathbf{E}}[\texttt{csuccess}|\texttt{hgpa}, \texttt{sat}, \texttt{iexam}] \\
&\quad = .70\texttt{hgpa} + 1.01(\texttt{sat}+1) + 0.18\texttt{iexam} + 5.45\texttt{iexam}/\texttt{hgpa}^2 - 1.43 \\
&\qquad - \left[.70\texttt{hgpa} + 1.01\texttt{sat} + 0.18\texttt{iexam} + 5.45\texttt{iexam}/\texttt{hgpa}^2 - 1.43\right] \\
&\quad = 1.01
\end{aligned}$$

- The estimated conditional-on-covariate effect of a 100-point increase in sat is a constant
- The conditional-on-covariate effect is the same as the population-averaged effect, because the conditional-on-covariate effect is a constant

Effect of a 10-point increase in iexam

Because `iexam` is measured in tens of points, the conditional-on-covarite effect of a 10-point increase in the `iexam` is estimated to be

$$\widehat{\mathbf{E}}[\texttt{csuccess}|\texttt{hgpa}, \texttt{sat}, (\texttt{iexam}+1)] - \widehat{\mathbf{E}}[\texttt{csuccess}|\texttt{hgpa}, \texttt{sat}, \texttt{iexam}]$$
$$= .70\texttt{hgpa} + 1.01\texttt{sat} + 0.18(\texttt{iexam}+1) + 5.45(\texttt{iexam}+1)/(\texttt{hgpa}^2) - 1.43$$
$$- \left[.70\texttt{hgpa} + 1.01\texttt{sat} + 0.18\texttt{iexam} + 5.45\texttt{iexam})/(\texttt{hgpa}^2) - 1.43\right]$$
$$= .18 + 5.45/\texttt{hgpa}^2$$

- The conditional-on-covariate effect varies with a student's high-school grade-point average
- The conditional-on-covariate effect differs from the population-averaged effect

What conditional-on-covariate effects tell us

- Suppose that I am a counselor who believes that only increases of .7 or more in `csuccess` matter
- A student with an `hgpa` of 4.0 asks me if a 10-point increase on the `iexam` will significantly affect his or her college success

```
. margins , expression(_b[iexam] + _b[it]/(hgpa^2)) at(hgpa=4)
Warning: expression() does not contain predict() or xb().
Predictive margins                              Number of obs     =      1,000
Model VCE    : Robust
Expression   : _b[iexam] + _b[it]/(hgpa^2)
at           : hgpa            =           4

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |     .51859    .0621809     8.34   0.000     .3967176    .6404623
------------------------------------------------------------------------------
```

- I tell the student “probably not”

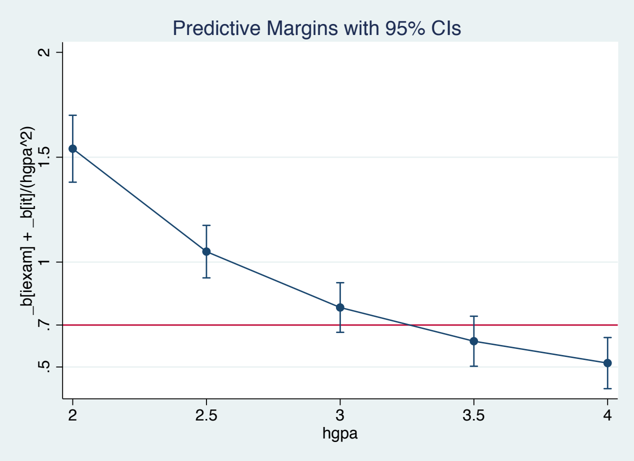
After the student leaves, I estimate the effect of a 10-point increase in iexam when hgpa is 2, 2.5, 3, 3.5, and 4

```
. margins , expression(_b[iexam] + _b[it]/(hgpa^2)) at(hgpa=(2 2.5 3 3.5 4))
Warning: expression() does not contain predict() or xb().
Predictive margins                              Number of obs     =      1,000
Model VCE    : Robust
Expression   : _b[iexam] + _b[it]/(hgpa^2)
1._at        : hgpa            =           2
2._at        : hgpa            =         2.5
3._at        : hgpa            =           3
4._at        : hgpa            =         3.5
5._at        : hgpa            =           4
```

	Margin	Delta-method Std. Err.	z	P>|z|	[95% Conf.	Interval]
_at						
1	1.5405	.0813648	18.93	0.000	1.381028	1.699972
2	1.049983	.0638473	16.45	0.000	.9248449	1.175122
3	.7835297	.0603343	12.99	0.000	.6652765	.9017828
4	.6228665	.0608185	10.24	0.000	.5036645	.7420685
5	.51859	.0621809	8.34	0.000	.3967176	.6404623

marginsplot

```
. quietly margins , expression(_b[iexam] + _b[it]/(hgpa^2)) ///
>         at(hgpa=(2 2.5 3 3.5 4))
. marginsplot , yline(.7) ylabel(.5 .7 1 1.5 2)
  Variables that uniquely identify margins: hgpa
```



Conditional-on-covariate inference

- Suppose $\mathbf{E}[y|x,\mathbf{z}]$ is my regression model for the outcome y as a function of x, whose effect I want to estimate, and $\mathbf{z}$, which are other variables on which I condition
- The regression function $\mathbf{E}[y|x,\mathbf{z}]$ tells me the mean of y for given values of x and $\mathbf{z}$
- The difference between the mean of y given x_1 and $\mathbf{z}$ and the mean of y given x_0 and $\mathbf{z}$ is an effect of x, and it is given by $\mathbf{E}[y|x=x_1,\mathbf{z}] - \mathbf{E}[y|x=x_0,\mathbf{z}]$
- This effect can vary with $\mathbf{z}$; it might be scientifically and statistically significant for some values of $\mathbf{z}$ and not for others
- Doctors, consultants, and counselors want to know what these effects for specified covariate values.

Stata workflow

- Under the usual assumptions of correct specification, I estimate the parameters of $\mathbf{E}[y|x, \mathbf{z}]$ using `regress` or another command
- I then use `margins` and `marginsplot` to estimate effects of x
- I also frequently use `lincom`, `nlcom`, and `predictnl` to estimate effects of x for given $\mathbf{z}$ values.

Who cares about the population?

- Now, suppose that I am a university administrator who believes that assigning enough tutors to the course will raise each student's `iexam` score by 10 points
  - I need a single measure that accounts for the distribution of the effects over individual students
- I use `margins` to estimate the mean college-success score that is observed when each student gets his or her current `iexam` score and to estimate the mean college-success score that would be observed when each student gets an extra 10 points on his or her `iexam` score.

Margins also estimates population-averaged effects

```
. margins , at(iexam = generate(iexam))   ///
>          at(iexam = generate(iexam+1) it = generate((iexam+1)/(hgpa^2)))
Predictive margins                              Number of obs     =      1,000
Model VCE    : Robust
Expression   : Linear prediction, predict()
1._at        : iexam           = iexam
2._at        : iexam           = iexam+1
               it              = (iexam+1)/(hgpa^2)
```

	Margin	Delta-method Std. Err.	t	P>|t|	[95% Conf.	Interval]
_at						
1	20.76273	.0434416	477.95	0.000	20.67748	20.84798
2	21.48141	.0744306	288.61	0.000	21.33535	21.62747

- 1._at estimates the mean college-success score when each student gets his or her current iexam score
- 2._at estimates the mean college-success score when each student gets an extra 10 points on his or her iexam score

- The average of the predicted values when each student gets his or her current `iexam` score, `yhat0`, matches the estimate reported by **margins** for `_at.1`
- The average of the predicted values when each student gets an extra 10 points on his or her `iexam` score, `yhat1`, matches the estimate reported by **margins** for `_at.2`

```
. preserve
. predict double yhat0
(option xb assumed; fitted values)
. replace iexam       = iexam + 1
(1,000 real changes made)
. replace it          = (iexam)/(hgpa^2)
(1,000 real changes made)
. predict double yhat1
(option xb assumed; fitted values)
. summarize yhat0 yhat1
```

Variable	Obs	Mean	Std. Dev.	Min	Max
yhat0	1,000	20.76273	1.625351	17.33157	26.56351
yhat1	1,000	21.48141	1.798292	17.82295	27.76324

```
. restore
```

Look at contrast option for margins

- Now, I use the `contrast` option to estimate the difference between the average of `csuccess` when each student gets an extra 10 points and the average of `csuccess` when each student gets his or her original score

```
. margins , at(iexam = generate(iexam))   ///
>         at(iexam = generate(iexam+1) it = generate((iexam+1)/(hgpa^2))) ///
>         contrast(atcontrast(r._at) nowald)
Contrasts of predictive margins
Model VCE     : Robust
Expression    : Linear prediction, predict()
1._at         : iexam            = iexam
2._at         : iexam            = iexam+1
                it               = (iexam+1)/(hgpa^2)
```

	Contrast	Delta-method Std. Err.	[95% Conf.	Interval]
_at (2 vs 1)	.7186786	.0602891	.6003702	.836987

- The "Delta-method" standard error takes the covariate observations as fixed and accounts only for the parameter estimation error
  Sample treatment effect for this particular batch of students
- The option `vce(unconditional)` gets me inference for the population from which I can repeatedly draw samples of students (Population treatment effect)

```
. margins , at(iexam = generate(iexam))   ///
>         at(iexam = generate(iexam+1) it = generate((iexam+1)/(hgpa^2))) ///
>         contrast(atcontrast(r._at) nowald) vce(unconditional)
Contrasts of predictive margins
Expression    : Linear prediction, predict()
1._at         : iexam             = iexam
2._at         : iexam             = iexam+1
                it                = (iexam+1)/(hgpa^2)
```

	Contrast	Unconditional Std. Err.	[95% Conf.	Interval]
_at (2 vs 1)	.7186786	.0609148	.5991425	.8382148

The difference in means is the mean of differences

- Suppose $\mathbf{E}[y|x, \mathbf{z}]$ is my regression model for the outcome y as a function of x, whose effect I want to estimate, and $\mathbf{z}$, which are other variables on which I condition
- The difference between the mean of y given x_1 and the mean of y given x_0 is an effect of x that has been averaged over the distribution of $\mathbf{z}$,

$$\begin{aligned} &\mathbf{E}[y|x = x_1] - \mathbf{E}[y|x = x_0] \\ &\quad = \mathbf{E}_{\mathbf{Z}}\left[\mathbf{E}[y|x = x_1, \mathbf{z}]\right] - \mathbf{E}_{\mathbf{Z}}\left[\mathbf{E}[y|x = x_0, \mathbf{z}]\right] \\ &\quad = \mathbf{E}_{\mathbf{Z}}\left[\mathbf{E}[y|x = x_1, \mathbf{z}] - \mathbf{E}[y|x = x_0, \mathbf{z}]\right] \end{aligned}$$

- The difference in the means that condition only the hypothesized x values is the mean of the diffences that condition on x and z
- The difference in the marginal effects is the mean of the conditional effects

Representative sample need apply

- Under the usual assumptions of correct specification, I can estimate the parameters of $\mathbf{E}[y|x, \mathbf{z}]$ using `regress` or another command
- I can then use `margins` and `marginsplot` to estimate a mean of these effects of x
  - The sample must be representative, perhaps after weighting, in order for the estimated mean of the effects to converge to a population mean.

```
. logit graduate hgpa sat iexam it, nolog
Logistic regression                             Number of obs     =      1,000
                                                LR chi2(4)        =     576.12
                                                Prob > chi2       =     0.0000
Log likelihood = -404.75078                     Pseudo R2         =     0.4158
```

graduate	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
hgpa	2.347051	.3975215	5.90	0.000	1.567923	3.126178
sat	1.790551	.1353122	13.23	0.000	1.525344	2.055758
iexam	1.447134	.1322484	10.94	0.000	1.187932	1.706336
it	1.713286	.7261668	2.36	0.018	.2900249	3.136546
_cons	-46.82946	3.168635	-14.78	0.000	-53.03987	-40.61905

The estimates imply that

$$\widehat{\texttt{Pr}}[\texttt{graduate} = 1|\texttt{hgpa}, \texttt{sat}, \texttt{iexam}]$$
$$= \texttt{F}\left[2.35\texttt{hgpa} + 1.79\texttt{sat} + 1.45\texttt{iexam} + 1.71\texttt{iexam}/(\texttt{hgpa}^2) - 46.83\right]$$

where $\texttt{F}(\mathbf{x}\boldsymbol{\beta}) = \exp(\mathbf{x}\boldsymbol{\beta})/[1 + \exp(\mathbf{x}\boldsymbol{\beta})]$ is the logistic distribution and $\widehat{\texttt{Pr}}[\texttt{graduate} = 1|\texttt{hgpa}, \texttt{sat}, \texttt{iexam}]$ denotes the estimated conditional probability function.

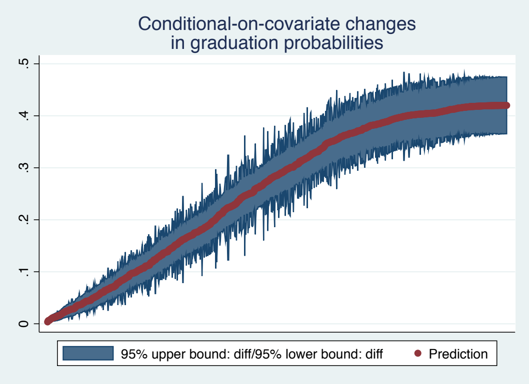
- Suppose that I am a researcher who wants to know the conditional-on-covariate effect of getting a 1400 instead of a 1300 on the SAT on the conditional graduation probability
- Because sat is measured in hundreds of points, the effect is estimated to be

$$
\begin{aligned}
&\widehat{\texttt{Pr}}[\texttt{graduate} = 1|\texttt{sat} = 14, \texttt{hgpa}, \texttt{iexam}] \\
&\qquad - \widehat{\texttt{Pr}}[\texttt{graduate} = 1|\texttt{sat} = 13, \texttt{hgpa}, \texttt{iexam}] \\
&= \texttt{F}\left[2.35\texttt{hgpa} + 1.79(14) + 1.45\texttt{iexam} + 1.71\texttt{iexam}/(\texttt{hgpa}^2) - 46.83\right] \\
&- \texttt{F}\left[2.35\texttt{hgpa} + 1.79(13) + 1.45\texttt{iexam} + 1.71\texttt{iexam}/(\texttt{hgpa}^2) - 46.83\right]
\end{aligned}
$$

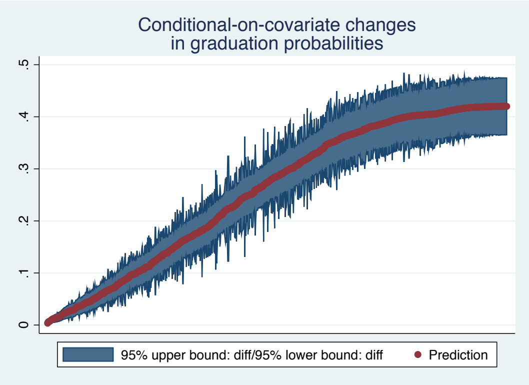
- The estimated conditional-on-covariate effect of going from 1300 to 1400 on the SAT varies over the values of hgpa and iexam, because F() is nonlinear

- I use `predictnl` to estimate these effects for each observation in the sample and then I graph them

```
. predictnl double diff =
>       ///
>    logistic( _b[hgpa]*hgpa + _b[sat]*14 + _b[iexam]*iexam + _b[it]*it + _b[_c
> ons]) ///
>  - logistic( _b[hgpa]*hgpa + _b[sat]*13 + _b[iexam]*iexam + _b[it]*it + _b[_c
> ons]) ///
>     , ci(low up)
note: confidence intervals calculated using Z critical values

. sort diff

. generate ob = _n

. twoway (rarea up low ob) (scatter diff ob) , xlabels(none) xtitle("") ///
>    title("Conditional-on-covariate changes" "in graduation probabilities")
```

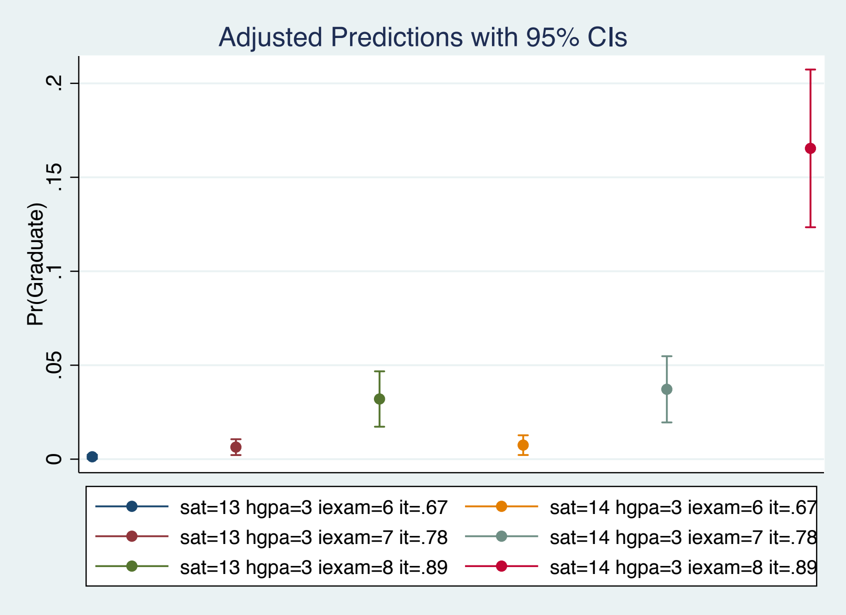


- the estimated differences in conditional graduation probabilities caused by going from 1300 to 1400 on the SAT range from close to 0 to more than .4 over the sample values of `hgpa` and `iexam`



- If I were a counselor advising specific students on the basis of their `hgpa` and `iexam` values
  - I would be interested in which students had effects near zero and in which students had effects greater than, say, .3
  - Methodologically, I would be interested in effects conditional on the covariates `hgpa` and `iexam`

```
. margins , at(sat=13 hgpa=3 iexam=6 it=.67)      ///
>           at(sat=13 hgpa=3 iexam=7 it=.78)      ///
>           at(sat=13 hgpa=3 iexam=8 it=.89)      ///
>           at(sat=14 hgpa=3 iexam=6 it=.67)      ///
>           at(sat=14 hgpa=3 iexam=7 it=.78)      ///
>           at(sat=14 hgpa=3 iexam=8 it=.89)      ///
>           noatlegend
Adjusted predictions                            Number of obs     =      1,000
Model VCE     : OIM

Expression    : Pr(graduate), predict()

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         _at |
          1  |   .0012537   .0005577     2.25   0.025     .0001605    .0023468
          2  |   .0064013   .0021517     2.97   0.003      .002184    .0106186
          3  |   .0320079    .007524     4.25   0.000     .0172612    .0467546
          4  |   .0074661   .0026775     2.79   0.005     .0022183     .012714
          5  |   .0371732   .0089876     4.14   0.000     .0195578    .0547885
          6  |   .1653855   .0214073     7.73   0.000     .1234281     .207343
------------------------------------------------------------------------------

. marginsplot, plotdim(_atopt) xdim(_atopt)       ///
>     xtitle("") xlabel(none)                     ///
>     legend(size(*.93) colfirst)
  Variables that uniquely identify margins: _atopt
  Multiple at() options specified:
      _atoption=1: sat=13 hgpa=3 iexam=6 it=.67
      _atoption=2: sat=13 hgpa=3 iexam=7 it=.78
      _atoption=3: sat=13 hgpa=3 iexam=8 it=.89
```


Adjusted Predictions with 95% CIs
Pr(Graduate)
0
.05
.1
.15
.2
sat=13 hgpa=3 iexam=6 it=.67
sat=13 hgpa=3 iexam=7 it=.78
sat=13 hgpa=3 iexam=8 it=.89
sat=14 hgpa=3 iexam=6 it=.67
sat=14 hgpa=3 iexam=7 it=.78
sat=14 hgpa=3 iexam=8 it=.89

- Now suppose I want to know "whether going from 1300 to 1400 on the SAT matters"
- I am thus interested in a single aggregate measure
- I use `margins` to estimate the mean of the conditional-on-covariate effects

```
. margins , at(sat=(13 14)) contrast(atcontrast(r._at) nowald)
Contrasts of predictive margins
Model VCE      : OIM
Expression     : Pr(graduate), predict()
1._at          : sat              =          13
2._at          : sat              =          14
```

	Contrast	Delta-method Std. Err.	[95% Conf.	Interval]
_at (2 vs 1)	.2576894	.0143522	.2295597	.2858192

- The difference in the mean graduation probabilities caused by going from 1300 to 1400 on the SAT is estimated to be .26

- The mean change is the same as the difference in the probabilities that are only conditioned on the hypothesized sat values

$$
\begin{aligned}
&\mathbf{E}\left[\widehat{\Pr}[\texttt{graduate}=1|\texttt{sat}=14,\texttt{hgpa},\texttt{iexam}]\right.\\
&\quad\left.-\widehat{\Pr}[\texttt{graduate}=1|\texttt{sat}=13,\texttt{hgpa},\texttt{iexam}]\right]\\
&=\mathbf{E}\left[\widehat{\Pr}[\texttt{graduate}=1|\texttt{sat}=14,\texttt{hgpa},\texttt{iexam}]\right]\\
&\quad-\mathbf{E}\left[\widehat{\Pr}[\texttt{graduate}=1|\texttt{sat}=13,\texttt{hgpa},\texttt{iexam}]\right]\\
&=\widehat{\Pr}[\texttt{graduate}=1|\texttt{sat}=14]-\widehat{\Pr}[\texttt{graduate}=1|\texttt{sat}=13]
\end{aligned}
$$

- The mean of the differences in the conditional probabilities is a difference in marginal probabilities
- The difference in the probabilities that condition only the values that define the "treatment" values is one of the population parameters that a potential-outcome approach would specify to be of interest

Odds ratio

- The odds of an event specifies how likely it is to occur, with higher values implying that the event is more likely
- An odds ratio is the ratio of the odds of an event in one scenario to the odds of the same event under a different scenario
- I am interested in the ratio of the graduation odds when a student has an SAT of 1400 to the graduation odds when a student has an SAT of 1300
- A value greater than 1 implies that going from 1300 to 1400 has raised the graduation odds
- A value less than 1 implies that going from 1300 to 1400 has lowered the graduation odds.

- The logistic model for the conditional probability implies that the ratio of the odds of graduation conditional on `sat`=14, `hgpa`, and `iexam` to the odds of graduation conditional on `sat`=13, `hgpa`, and `iexam` is exp(_b[sat])

```
. logit , or
Logistic regression                             Number of obs     =      1,000
                                                LR chi2(4)        =     576.12
                                                Prob > chi2       =     0.0000
Log likelihood = -404.75078                     Pseudo R2         =     0.4158
```

graduate	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
hgpa	10.45469	4.155964	5.90	0.000	4.796674	22.78673
sat	5.992756	.8108931	13.23	0.000	4.596726	7.812761
iexam	4.250916	.5621767	10.94	0.000	3.280292	5.508743
it	5.547158	4.028162	2.36	0.018	1.336461	23.02421
_cons	4.59e-21	1.46e-20	-14.78	0.000	9.23e-24	2.29e-18

- The conditional-on-covariate graduation odds are estimated to be 6 times higher for a student with a 1400 SAT than for a student with a 1300 SAT

This result comes from some algebra that shows that

$$\frac{\frac{\widehat{\Pr}[\texttt{graduate}=1|\texttt{sat}=14,\texttt{hgpa},\texttt{iexam}]}{1-\widehat{\Pr}[\texttt{graduate}=1|\texttt{sat}=14,\texttt{hgpa},\texttt{iexam}]}}{\frac{\widehat{\Pr}[\texttt{graduate}=1|\texttt{sat}=13,\texttt{hgpa},\texttt{iexam}]}{1-\widehat{\Pr}[\texttt{graduate}=1|\texttt{sat}=13,\texttt{hgpa},\texttt{iexam}]}} = \exp\left(\_\texttt{b}[\texttt{sat}]\right)$$

when

$$\widehat{\Pr}[\texttt{graduate} = 1|\texttt{sat}, \texttt{hgpa}, \texttt{iexam}] = \frac{\exp(\mathbf{x}\boldsymbol{\beta})}{1+\exp(\mathbf{x}\boldsymbol{\beta})}$$

where $\mathbf{x}\boldsymbol{\beta} =$
$\_\texttt{b}[\texttt{hgpa}]\texttt{hgpa} + \_\texttt{b}[\texttt{sat}]\texttt{sat} + \_\texttt{b}[\texttt{iexam}]\texttt{iexam} + \_\texttt{b}[\texttt{it}]\texttt{it} + \_\texttt{b}[\_\texttt{cons}]$

- More generally, `exp(_b[sat])` is the ratio of the conditional-on-covariate graduation odds for a student getting one more unit of `sat` to the conditional-on-covariate graduation odds for a student getting his or her current `sat` value

Two highlights

- I want to highlight that
  - the logistic functional form makes this conditional-on-covariate odds ratio a constant
  - the ratio of conditional-on-covariate odds differs from the ratio of odds that condition only the hypothesized values

Computing a conditional-on-covariate odds ratio

- the conditional-on-covariate odds ratio does not vary over the covariate patterns in the sample

```
. generate sat_orig = sat
. replace sat  = 13
(999 real changes made)
. predict double pr0
(option pr assumed; Pr(graduate))
. replace sat  = 14
(1,000 real changes made)
. predict double pr1
(option pr assumed; Pr(graduate))
. replace sat  = sat_orig
(993 real changes made)
. generate orc = (pr1/(1-pr1))/(pr0/(1-pr0))
. summarize orc
```

Variable	Obs	Mean	Std. Dev.	Min	Max
orc	1,000	5.992756	0	5.992756	5.992756

That the standard deviation is 0 highlights that the values are constant.

Conditional-on-hypothesized-values-only odds ratio

- Use margins to estimate the ratio of graduation odds that condition only on the hypothesized sat values

```
. margins , at(sat=(13 14)) post
Predictive margins                              Number of obs     =      1,000
Model VCE     : OIM
Expression    : Pr(graduate), predict()
1._at         : sat             =          13
2._at         : sat             =          14

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         _at |
          1  |   .2430499    .018038    13.47   0.000     .2076961    .2784036
          2  |   .5007393   .0133553    37.49   0.000     .4745634    .5269152
------------------------------------------------------------------------------

. nlcom (_b[2._at]/(1-_b[2._at]))/(_b[1._at]/(1-_b[1._at]))
       _nl_1:  (_b[2._at]/(1-_b[2._at]))/(_b[1._at]/(1-_b[1._at]))

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |   3.123606   .2418127    12.92   0.000     2.649661     3.59755
------------------------------------------------------------------------------
```

Defining a conditional-on-hypothesized-values-only odds ratio

- Mathematically, this estimate implies that

$$\frac{\frac{\widehat{\texttt{Pr}}[\texttt{graduate}=1|\texttt{sat}=14]}{1-\widehat{\texttt{Pr}}[\texttt{graduate}=1|\texttt{sat}=14]}}{\frac{\widehat{\texttt{Pr}}[\texttt{graduate}=1|\texttt{sat}=13]}{1-\widehat{\texttt{Pr}}[\texttt{graduate}=1|\texttt{sat}=13]}} = 3.12$$

- The Delta-method standard error provides inference for the students in this sample as opposed to an unconditional standard error that provides inference for repeated sample from the population

Why they differ

- The mean of a nonlinear function differs from a nonlinear function evaluated at the mean

$$\frac{\frac{\Pr[\text{graduate}=1|\text{sat}=14]}{1-\Pr[\text{graduate}=1|\text{sat}=14]}}{\frac{\Pr[\text{graduate}=1|\text{sat}=13]}{1-\Pr[\text{graduate}=1|\text{sat}=13]}} = \frac{\frac{\mathbf{E}[\Pr[\text{graduate}=1|\text{sat}=14,\text{hgpa},\text{iexam}]]}{1-\mathbf{E}[\Pr[\text{graduate}=1|\text{sat}=14,\text{hgpa},\text{iexam}]]}}{\frac{\mathbf{E}[\Pr[\text{graduate}=1|\text{sat}=13,\text{hgpa},\text{iexam}]]}{1-\mathbf{E}[\Pr[\text{graduate}=1|\text{sat}=13,\text{hgpa},\text{iexam}]]}}$$

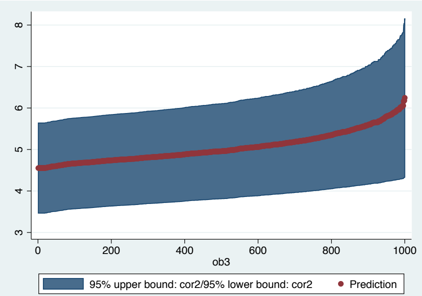
$$\neq \mathbf{E}\left[\frac{\frac{\Pr[\text{graduate}=1|\text{sat}=14,\text{hgpa},\text{iexam}]}{1-\Pr[\text{graduate}=1|\text{sat}=14,\text{hgpa},\text{iexam}]}}{\frac{\Pr[\text{graduate}=1|\text{sat}=13,\text{hgpa},\text{iexam}]}{1-\Pr[\text{graduate}=1|\text{sat}=13,\text{hgpa},\text{iexam}]}}\right]$$

Which one do want?

- Which odds ratio is of interest depends on what you want to know
    - The conditional-on-covariate odds ratio is of interest when conditional-on-covariate comparisons are the goal
    - The ratio of the odds that condition only on hypothesized `sat` values is the population parameter that a potential-outcome approach would specify to be of interest

- I use `predictnl` to compute conditional-on-covariate odds ratios of going from a 70 to an 80 on the short-course exam `iexam`

```
. local xb1 " _b[hgpa]*hgpa + _b[sat]*sat + _b[iexam]*8 + _b[it]*(8/hgpa^2) + _
> b[_cons]"
. local pr1 "logistic(`xb1´)"
. local xb0 " _b[hgpa]*hgpa + _b[sat]*sat + _b[iexam]*7 + _b[it]*(7/hgpa^2) + _
> b[_cons]"
. local pr0 "logistic(`xb0´)"
. predictnl double cor2 = (`pr1´/(1-`pr1´))/(`pr0´/(1-`pr0´)), ci(low2 up2)
note: confidence intervals calculated using Z critical values
. sort cor2
. generate ob3 = _n
. twoway (rarea up2 low2 ob3) (scatter cor2 ob3)
```



- Use `margins` to estimate the ratio of graduation odds that condition only on the hypothesized `iexam` values

```
. margins , at(iexam=7 it=generate(7/(hgpa^2)))          ///
>           at(iexam=8 it=generate(8/(hgpa^2))) post
Predictive margins                              Number of obs     =      1,000
Model VCE    : OIM

Expression   : Pr(graduate), predict()

1._at        : iexam           =           7
               it              = 7/(hgpa^2)

2._at        : iexam           =           8
               it              = 8/(hgpa^2)

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         _at |
          1  |   .1797364   .0170616    10.53   0.000     .1462962    .2131766
          2  |   .3711477   .0150742    24.62   0.000     .3416028    .4006926
------------------------------------------------------------------------------

. nlcom (_b[2._at]/(1-_b[2._at]))/(_b[1._at]/(1-_b[1._at]))
       _nl_1:  (_b[2._at]/(1-_b[2._at]))/(_b[1._at]/(1-_b[1._at]))

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _nl_1 |   2.693491    .178482    15.09   0.000     2.343673    3.043309
------------------------------------------------------------------------------
```

Cameron, A. C., and P. K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. Cambridge: Cambridge University Press.

Imbens, G. W. 2004. Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and statistics* 86(1): 4–29.

Imbens, G. W., and J. M. Wooldridge. 2009. Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature* 47: 5–86.

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