Lecture 6: Epi commands and tables

BIOSTAT 212, Summer 2018

UCSF

Kristen Aiemjoy
Department of Epidemiology and Biostatistics
University of California San Francisco

Displaying exact p-value after a linear regression

Option 1

```
display 2*ttail(df, abs(t))
```

Option 2

Use "`est tab`" command after regression

```
regress var1 var2
est tab, p(%-5.1e)
```

Set stata to display more digits for p-value

```
set pformat %-5.1e
```

Set permanently:

```
set pformat %-5.1e, permanently
```

Go back to 4 decimal places:

```
set pformat %5.4f
```

Outline for today

- **Introduction to basic epidemiology concepts**
- Estimating Measures of Association in Stata
  - Risk ratios & differences
  - Rate ratios & differences
  - Odds ratios
  - Logistic regression
  - Binary regression
- Immediate commands for epi
- Creating and exporting tables
  - `table1`
  - `outreg2`
  - `tabout`

Load NHANES data

```
webuse nhanes2
```

Exposures and Outcomes

- Does “exposure” cause “outcome”?

$$\text{High BP} \xrightarrow[?]{} \text{Heart attack}$$

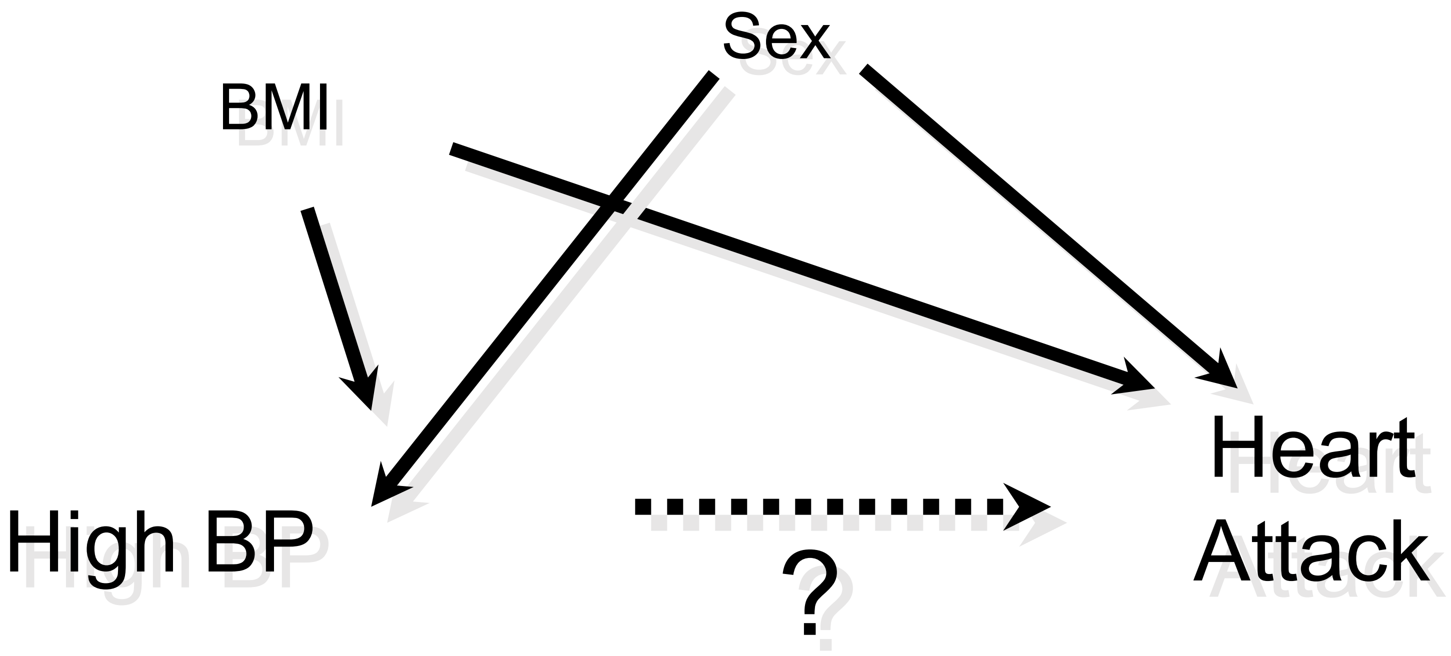
Causal Inference

6 ways a statistical association can occur:

1. Chance
2. Selection bias
3. Measurement bias
4. Confounding
5. Reverse causation
6. **True causal relationship**

Confounding

- Do other variables mask or exaggerate the association between exposure and outcome?



Outline for today

- Introduction to basic epidemiology concepts
- **Estimating Measures of Association in Stata**
  - Risk ratios & differences
  - Rate ratios & differences
  - Odds ratios
  - Logistic regression
  - Binary regression
- Immediate commands for epi
- Creating and exporting tables
  - `table1`
  - `outreg2`
  - `tabout`

Measures of Association

- An association between an exposure and outcome can be assessed by comparing the probability of outcome in the exposed group to the probability of outcome in unexposed group.

- **Ratio (division)**

  $$\frac{\text{[Pr(outcome) in exposed]}}{\text{[Pr(outcome) in unexposed]}}$$

- **Difference (subtraction)**

  [Pr(outcome) in exposed] – [Pr(outcome) in unexposed]

Ratio Measures of Association

Study Design	Ratio Measure	Stata command

Cross-sectional study	Prevalence Ratio (PR)	`cs O E, by(C)`
		`binreg O E C, rr`
		`glm O E C, family(bin) link(log) eform`

Cohort study	Risk Ratio (RR)	`cs O E, by(C)`
		`binreg O E C, rr`
		`glm O E C, family(bin) link(log) eform`

	Rate Ratio (RR)	`ir O E T` `poisson O E C, irr`
	Hazard Ratio (HR)	`stcox O E C`

Case-control study	Odds Ratio (OR)	`cc O E, by(C)`
		`mhodds O E, by(C)`
		`logistic O E C, or`
		`tabodds O E, adjust(C)`

O = outcomevar E = exposurevar C = covariates T = time

Difference Measures of Association

Study Design	Ratio Measure	Stata command
Cross-sectional study	Prevalence Difference	`cs O E` `binreg O E C, rd`
Cohort study	Risk Difference	`cs O E` `binreg O E C, rd`
	Rate Difference	`ir O E T, ird` `poisson O E C`
Case-control study	N/A	

`O = outcomevar E = exposurevar C = covariates T = time`

Outline for today

- Introduction to basic epidemiology concepts
- Estimating Measures of Association in Stata
  - **Risk ratios & differences**
  - Rate ratios & differences
  - Odds ratios
  - Logistic regression
  - Binary regression
- Immediate commands for epi
- Creating and exporting tables
  - `table1`
  - `outreg2`
  - `tabout`

Risk Ratio and Risk Difference

		Disease: Yes	Disease: No
Exposure	Yes	a	b
	No	c	d

Risk Ratio =

$$\frac{a/(a+b)}{c/(c+d)}$$

Risk Difference =

$$a/(a+b) - c/(c+d)$$

Risk = Probability

cohort studies & RCTs

*similar concept for prevalence ratios, used in cross-sectional studies

Risk Ratios/Differences: `cs`

`cs outcomevar predictorvar`

– calculates risk ratios and risk differences

<u>Options</u>

`,by (varname)`	-stratify by a third variable
`[, missing]`	-include missing data as a category of the stratified variable

<u>Limitations</u>

- Cannot adjust for more than 1 covariate

Practice

- What is the RR for the effect of high BP on heart attack?

```
cs heartatk highbp
di (279/4376)/(197/5973)
```

- What is the risk difference?

```
cs heartatk highbp
di (279/4376)-(197/5973)
```

Outline for today

- Introduction to basic epidemiology concepts
- Estimating Measures of Association in Stata
  - Risk ratios & differences
  - **Rate ratios & differences**
  - Odds ratios
  - Logistic regression
  - Binary regression
- Immediate commands for epi
- Creating and exporting tables
  - `table1`
  - `outreg2`
  - `tabout`

Rate Ratio and Rate Difference

Rate Ratio =

$$\frac{\text{[Cases in exposed/exposed person-time]}}{\text{[Cases in unexposed/unexposed person-time]}}$$

Rate Difference =

[Cases in exposed/exposed person-time] -[Cases in unexposed/unexposed person-time]

Use:

```
ir O E T
poisson O E C, irr
cox O E C
```

```
O = outcomevar
E = exposurevar
C = covariates
T = time
```

For cohort studies

Outline for today

- Introduction to basic epidemiology concepts
- Estimating Measures of Association in Stata
  - Risk ratios & differences
  - Rate ratios & differences
  - **Odds ratios**
  - Logistic regression
  - Binary regression
- Immediate commands for epi
- Creating and exporting tables
  - `table1`
  - `outreg2`
  - `tabout`

Odds Ratio

Outcome

Exposure		Yes	No
	Yes	a	b
	No	c	d

Odds Ratio =

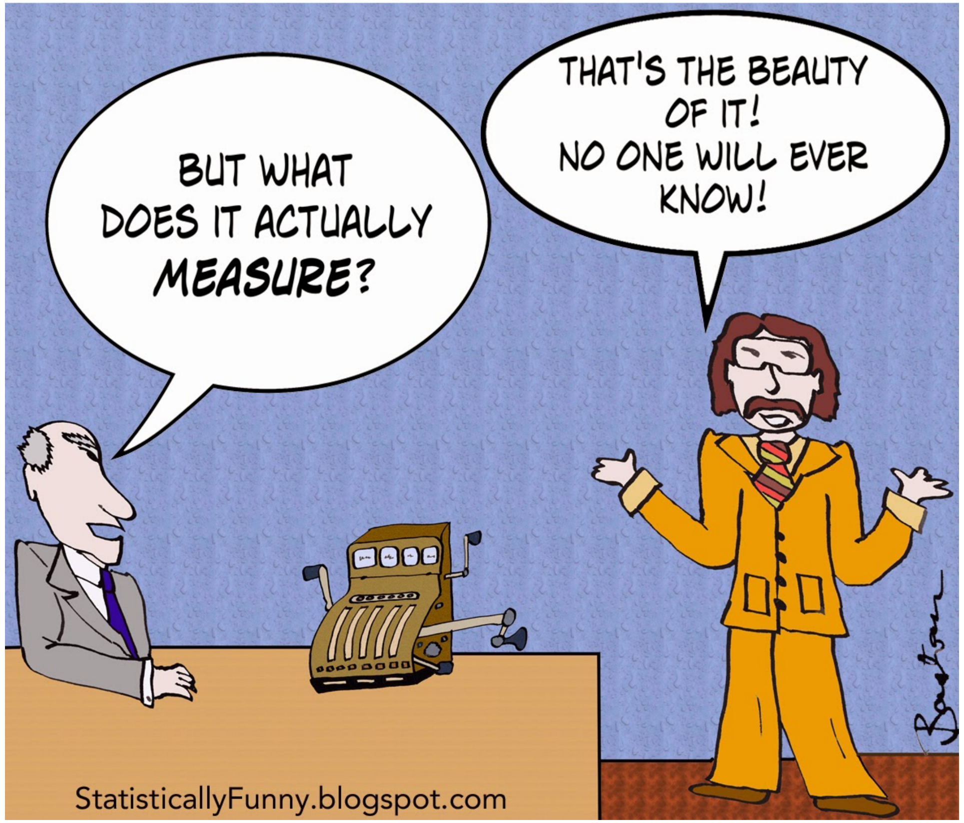
$$\frac{a/b}{c/d}$$

For cross-control studies

The odd thing about Odds Ratios….

- Odds are not-intuitive, difficult to communicate and interpret

- They are farther from the null than the risk ratio
  - exaggerate the association if interpreted as a probability



Odds Ratios: `cc`

`cc outcomevar predictorvar`

– calculates odds ratios

Options

`,by (varname)`	-stratify by a third variable
`[, missing]`	-include missing data as a category of the stratified variable

Limitations

Cannot adjust for more than 1 covariate

Odds Ratios: `mhodds`

`mhodds outcomevar predictorvar [covariates]`

– calculates odds ratios and can adjust for covariates

Options

`,by (varname)` -stratify by a third variable [to look for interaction]

`[, missing]` -include missing data as a category of the stratified variable

Limitations

- Uses stratification to adjust for covariates
  - with several covariates, strata start to get "thin"

Practice

- What is the OR for the effect of high BP on heart attack?

```
cc heartatk highbp

mhodds heartatk highbp

cs heartatk highbp, or

di (279/4097)/(197/5776)
```

Outline for today

- Introduction to basic epidemiology concepts
- Estimating Measures of Association in Stata
  - Risk ratios & differences
  - Rate ratios & differences
  - Odds ratios
  - **Logistic regression**
  - Binary regression
- Immediate commands for epi
- Creating and exporting tables
  - `table1`
  - `outreg2`
  - `tabout`

Logistic regression

```
logistic outcomevar predictorvar covariates
```

```
logit outcomevar predictorvar covariates, or
```

```
glm outcomevar predictorvar covariates, family(binomial)
     link(logit) eform
```

<u>Options</u>

`##` -Specify an interaction between 2 variables

If using any categorical predictors with >2 categories, use `i.`*`catvar`*

Practice

- What is the odds ratio (OR) for the effect of high BP on heart attack?

```
logistic heartatk highbp
logit heartatk highbp, or
glm heartatk highbp, family(binomial) link(logit) eform
```

- What is the odds ratio (OR) adjusted for sex?

```
logistic heartatk highbp sex
logit heartatk highbp sex, or
glm heartatk highbp sex, family(binomial) link(logit) eform
mhodds heartatk highbp sex
```

Outline for today

- Introduction to basic epidemiology concepts
- Estimating Measures of Association in Stata
  - Risk ratios & differences
  - Rate ratios & differences
  - Odds ratios
  - Logistic regression
  - **Binary regression**
- Immediate commands for epi
- Creating and exporting tables
  - `table1`
  - `outreg2`
  - `tabout`

Binary regression

`binreg` fits generalized linear models for the binomial family. It estimates odds ratios, risk ratios, health ratios, and risk differences.

`binreg outcomevar predictorvar covariates`

<u>Options</u>

`, rr`	-risk ratio
`, rd`	-risk difference
`, or`	-odds ratio
`, hr`	-hazard ratio

Practice

- What is the risk ratio (OR) for the effect of high BP on heart attack adjusted for sex?

```
binreg heartatk highbp sex, rr
```

Outline for today

- Introduction to basic epidemiology concepts
- Estimating Measures of Association in Stata
  - Risk ratios & differences
  - Rate ratios & differences
  - Odds ratios
  - Logistic regression
  - Binary regression
- **Immediate commands for epi**
- Creating and exporting tables
  - `table1`
  - `outreg2`
  - `tabout`

Immediate commands: Epidemiology

- Examples:
  - `iri` incidence rate calculator
    - `iri #exposed_cases #unexposed_cases #exposed_time #unexposed_time`
  - `cci` odds ratio calculator (**c**ase-**c**ontrol study)
    - `cci #exposed_cases #unexposed_cases #exposed_controls #unexposed_controls`
  - `csi` risk/ risk ratio calculator (**c**ohort **s**tudy)
    - `cci #exposed_cases #unexposed_cases #exposed_not_cases #unexposed_not_cases`

Abstract

Background: Despite high morbidity and mortality, the laboratory diagnosis of gastrointestinal infections is largely neglected in tropical African settings. This study aims to apply the Luminex multiplex PCR assay for the diagnosis of gastrointestinal pathogens in rural Ghana to evaluate its usefulness as a routine method.

Methods: A case–control study was conducted at the Agogo Presbyterian Hospital in Ghana. Stool samples were collected from children below 6 years of age with (cases) and without (controls) diarrhoea. Samples were screened for 15 different diarrhoeal pathogens by the Luminex xTAG GPP assay and associations between diarrhoea and gastrointestinal infections and fractions attributable to diarrhea (AF) were determined.

Results: The Luminex PCR assay identified organisms in 96.6 % ($n = 428$) of 443 cases and in 92.5 % ($n = 221$) of 239 selected controls. A mean of 2.5 (standard deviation [SD]: ±1.3) and 2.3 (SD: ±1.3) organisms per sample were detected in cases and controls respectively. An association with diarrhoea was found for rotavirus (adjusted odds ratio [aOR] = 7.2; 95 % confidence interval [CI]: 2.9–18.1), norovirus (aOR = 2.7; 95 % CI: 1.4–5.3) and *Shigella* spp. (aOR = 1.7; 95 % CI: 1.2–2.4) with respective AFs of 12.5 % (95 % CI: 9.6–15.3), 7.9 % (95 % CI: 3.8–11.7) and 16.9 % (95 % CI: 6.9–25.9).

Conclusion: The high proportion of pathogen-positive stool samples with a high number of co-infections in cases and controls suggests a substantial amount of transient or colonizing microorganisms for which treatment is not necessarily implicated. The use of sequential diagnostic algorithms with pathogen specific or quantitative PCRs might be most appropriate for diagnosing gastrointestinal infections.

Keywords: Ghana, West Africa, Multiplex Polymerase Chain Reaction, Gastrointestinal infections

Eibach D, Krumkamp R, Hahn A, et al. Application of a multiplex PCR assay for the detection of gastrointestinal pathogens in a rural African setting. *BMC Infectious Diseases*. 2016;16:150. doi:10.1186/s12879-016-1481-7.

QUIZ

		Diarrhea	
		Cases	Controls
Rotavirus	+	64	5
	-	379	234

1. Calculate the odds ratio

______

```
cci #cases_(+)rotavirus #cases_(-)rotavirus #controls_(+)rotavirus #controls_(-)rotavirus
```

Eibach D, Krumkamp R, Hahn A, et al. Application of a multiplex PCR assay for the detection of gastrointestinal pathogens in a rural African setting. *BMC Infectious Diseases*. 2016;16:150. doi:10.1186/s12879-016-1481-7.

```
. cci 64 379 5 234
```

	Exposed	Unexposed	Total	Proportion Exposed	
Cases	64	379	443	0.1445	
Controls	5	234	239	0.0209	
Total	69	613	682	0.1012	
	Point estimate		[95% Conf. Interval]		
Odds ratio	7.902902		3.144012	25.4793	(exact)
Attr. frac. ex.	.8734642		.681935	.9607525	(exact)
Attr. frac. pop	.126189				

chi2(1) = 26.06 Pr>chi2 = 0.0000

Table 2 Crude odds ratio (OR), age-adjusted OR (aOR) on associations between detected pathogens and diarrhoea, sorted by number of cases

Pathogen[a]	Cases ($n = 443$)	Controls ($n = 239$)	Negative PCR-result	Crude OR	aOR
Escherichia coli LT/ST (ETEC)	247 (55.8)	144 (60.3)	291	0.8 (0.6–1.2)	0.9 (0.6–1.2)
Giardia lamblia	228 (51.5)	125 (52.3)	329	0.9 (0.7–1.3)	1.0 (0.7–1.3)
Shigella spp.	185 (41.8)	76 (32.6)	421	1.5 (1.1–2.2)	1.7 (1.2–2.4)
Campylobacter spp.	153 (34.5)	78 (32.6)	451	1.1 (0.8–1.5)	1.1 (0.8–1.6)
Rotavirus	64 (14.5)	5 (2.1)	613	7.9 (3.1–25.5)	7.2 (2.9–18.1)
Escherichia coli O157	59 (13.3)	37 (15.5)	586	0.7 (0.5–1.2)	0.8 (0.5–1.2)
Norovirus	55 (12.4)	11 (4.5)	616	2.9 (1.5–6.3)	2.7 (1.4–5.3)
Salmonella spp.	53 (12.0)	37 (15.5)	592	0.8 (0.5–1.3)	0.8 (0.5–1.3)
Cryptosporidium spp.	33 (7.5)	12 (5.0)	637	1.5 (0.7–3.3)	1.4 (0.7–2.9)
Adenovirus (41/41)	17 (3.8)	14 (5.9)	651	0.6 (0.3–1.4)	0.6 (0.3–1.3)
Escherichia coli STEC	9 (3.7)	12 (2.7)	661	0.7 (0.3–1.9)	0.7 (0.3–1.8)
Clostridium diffcile toxin A,B	2 (0.5)	2 (0.8)	678	-	-
Entamoaba histolytica	1 (0.2)	2 (0.8)	679	-	-

[a]*Vibrio cholerae* and *Yersinia enterocolitica* were not detected in any sample
OR odds ratio, *aOR* age adjusted odds ratio

Eibach D, Krumkamp R, Hahn A, et al. Application of a multiplex PCR assay for the detection of gastrointestinal pathogens in a rural African setting. *BMC Infectious Diseases*. 2016;16:150. doi:10.1186/s12879-016-1481-7.

Outline for today

- Introduction to basic epidemiology concepts
- Estimating Measures of Association in Stata
  - Risk ratios & differences
  - Rate ratios & differences
  - Odds ratios
  - Logistic regression
  - Binary regression
- Immediate commands for epi
- **Creating and exporting tables**
  - **`table1`**
  - `outreg2`
  - `tabout`

table1

```
ssc install table1
```

`table1` generates a "table 1" of characteristics for a manuscript, which may be either continuous or categorical. The observations are often grouped, with a "p-value" column on the right comparing the characteristics between groups.

```
table1, vars(var1 type \ var2 type)
```

<u>variable `type` options</u>

`contn` - continuous, normally distributed

`conts` - continuous, skew

`cat` - categorical, groups compared using Pearson's chi-squared

`cate` - categorical, groups compared using Fisher's exact test

`bin` - binary, groups compared using Pearson's chi-squared

`bine` - binary, groups compared using Fisher's exact test

table1, cont.

```
table1, by(varname) vars(var1 type \ var2 type)
```

<u>additional options</u>

`, by(varname)`	- report summary by a group variable
`test`	- include column describing the significance test
`pdp(#)`	- max number of decimal places in p-value
`onecol`	- report categorical variable levels underneath variable name instead of in separate column
`format(%fmt)`	- default display format for contn and conts variables
`cformat(%fmt)`	- default display format for cat and bin variables
`missing`	- don't exclude missing values from categorical variables
`saving(filename.xls)`	- save table to Excel file

Example

```
table1, by(heartatk) vars(bmi contn \ bpsystol contn \ sex cat) saving(sumtable.xls)
```

Factor	Level	heartatk = 0	heartatk = 1	p-value
N		9873	476	
Body Mass Index (BMI), mean (SD)		25.49638 (4.913941)	26.39891 (4.870273)	<0.001
systolic blood pressure, mean (SD)		130.4081 (23.23776)	140.813 (23.0926)	<0.001
1=male, 2=female				

A	B	C	D	E
Factor	Level	heartatk = 0	heartatk = 1	p-value
N		9873	476	
Body Mass Index (BMI), mean (SD)		25.49638 (4.913941)	26.39891 (4.870273)	<0.001
systolic blood pressure, mean (SD)		130.4081 (23.23776)	140.813 (23.0926)	<0.001
1=male, 2=female	Male	4597 (46.6%)	318 (66.8%)	<0.001
	Female	5276 (53.4%)	158 (33.2%)	

Outline for today

- Introduction to basic epidemiology concepts
- Estimating Measures of Association in Stata
  - Risk ratios & differences
  - Rate ratios & differences
  - Odds ratios
  - Logistic regression
  - Binary regression
- Immediate commands for epi
- Creating and exporting tables
  - `table1`
  - **`outreg2`**
  - `tabout`

outreg2

```
ssc install outreg2
```

- An add-on package to make nice looking regression output and summary tables
- export into .txt, word(.rtf), excel(.xml), LaTeX(.tex), or stata data file(.dta)
- Many options for selecting the statistics you want to include
- Easy to add title, notes
- Specify how you want numbers formatted (2 decimal points for example)

Example

```
logistic heartatk highbp sex
outreg2 using OR_Table, eform word
```

Will export a **.rtf** file you can open in word

File name, you choose. Will save in whatever working directory you are in myreg.rtf

VARIABLES	(1) posths_sev
posths_sev	
no_acu	1.810**
	(0.421)
health_bl	0.982***
	(0.00535)
Constant	1.514
	(0.535)
Observations	326

seEform in parentheses

*** p<0.01, ** p<0.05, * p<0.1

`outreg2` example ~ excel

```
logistic heartatk highbp sex
outreg2 using OR_Table, eform replace excel ctitle(odds
    ratio)
```

Replace last version with the same name, otherwise it will add-on as an additional column

Will export a **.xml** file you can open in excel

Change column title

myreg.xml

	A	B	C
1			
2		(1)	
3	VARIABLES	odds ratio	
4			
5	posths_sev		
6			
7	no_acu	1.810**	
8		(0.421)	
9	health_bl	0.982***	
10		(0.00535)	
11	Constant	1.514	
12		(0.535)	
13			
14	Observations	326	
15	seEform in parentheses		
16	*** p<0.01, ** p<0.05, * p<0.1		
17			

`outreg2` example ~ LaTeX

```
logistic heartatk highbp sex
outreg2 using OR_Table, eform replace tex(pretty)
ctitle(odds ratio)
```

Will export a **.tex** file you can open in your LaTeX writing program or on overleaf website

Formatting option which adds bolds and italics

Opening .tex files

1. Use a cloud-based reader
   https://www.overleaf.com
   (*Lots of cool templates*)

1. Download TeX and TeX Studio (Both free)
   https://www.latex-project.org/get/
   https://www.texstudio.org

2. R Studio can open and format .tex files (but you need to have TeX installed

1. Open with your text editor, copy and paste to overleaf.com
2. Open directly with TeX (if installed)

OR_Table.tex

myreg.tex

```
\documentclass[]{article}
\setlength{\pdfpagewidth}{8.5in} \setlength{\pdfpageheight}{11in}
\begin{document}
\begin{center}
\begin{tabular}{lc} \hline
 & (1) \\
VARIABLES & odds ratio \\ \hline
\vspace{4pt} & \begin{footnotesize}\end{footnotesize} \\
posths\_sev &  \\
\vspace{4pt} & \begin{footnotesize}\end{footnotesize} \\
no\_acu & 1.810** \\
\vspace{4pt} & \begin{footnotesize}(0.421)\end{footnotesize} \\
health\_bl & 0.982*** \\
\vspace{4pt} & \begin{footnotesize}(0.00535)\end{footnotesize} \\
Constant & 1.514 \\
 & \begin{footnotesize}(0.535)\end{footnotesize} \\
\vspace{4pt} & \begin{footnotesize}\end{footnotesize} \\
 Observations & 326 \\ \hline
\multicolumn{2}{c}{\begin{footnotesize} seEform in parentheses\end{footnotesize}} \\
\multicolumn{2}{c}{\begin{footnotesize} *** p$<$0.01, ** p$<$0.05, *
p$<$0.1\end{footnotesize}} \\
\end{tabular}
\end{center}
\end{document}
```

	(1)
VARIABLES	odds ratio
posths_sev	
no_acu	1.810**
	(0.421)
health_bl	0.982***
	(0.00535)
Constant	1.514
	(0.535)
Observations	326

seEform in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Overleaf demo

Free & Interactive Online Introduction to LaTeX

https://www.overleaf.com/latex/learn/free-online-introduction-to-latex-part-1#.WaTwo62ZNjQ

Outline for today

- Introduction to basic epidemiology concepts
- Estimating Measures of Association in Stata
  - Risk ratios & differences
  - Rate ratios & differences
  - Odds ratios
  - Logistic regression
  - Binary regression
- Immediate commands for epi
- Creating and exporting tables
  - `table1`
  - `outreg2`
  - **`tabout`**

tabout

```
ssc install tabout
```

- An add-on package to export summary tables
- Currently you can only export .txt or .csv files
- Beta version can export to docX and excel
  - http://tabout.net.au/docs/home.php
- Include several variables, summary statistics and test statistics of your choice

`tabout` example (acupuncture.dta)

```
tabout group using table1.csv, replace ///
cell(mean basehs mean oneyrhs mean posths) format(1p 1p 1p) ///
style(csv) sum clab("Baseline" "Post"  "1yrPost") npos(col)
```

	A	B	C	D	E
1	intervention group	Mean	Mean	Mean	N
2		Baseline	Post	1yrPost	N
3	control	27.50%	22.30%	24.50%	196
4	acupuncture	25.60%	16.20%	19.10%	205
5	Total	26.50%	19.10%	21.60%	401
6					

Additional table packages

```
ssc install sutex
```

sutex - Generates LaTeX code for summary statistics tables.

```
ssc install estout
```

estout - Regression tables

Extra slides

logistic annotated output

logistic heartatk highbp sex

```
Logistic regression                          Number of obs[a]   =     10,349
                                             LR chi2(2)[b]      =     119.17
                                             Prob > chi2[c]     =     0.0000
Log likelihood = -1871.0093[a]               Pseudo R2[d]       =     0.0309

------------------------------------------------------------------------------
    heartatk[f] | Odds Ratio[j]   Std. Err.[k]      z[l]    P>|z|[m]     [95% Conf. Interval][n]
-------------+----------------------------------------------------------------
      highbp |   1.872771[g]  .1794883     6.55   5.9e-11    1.552046    2.259772
         sex |   .4562092[h]  .0455626    -7.86   3.9e-15    .3751048    .5548498
       _cons |   .1082779[i]  .0168897   -14.25   4.4e-46    .0797563     .146999
------------------------------------------------------------------------------
Note: _cons estimates baseline odds.
```

Logistic regression annotated output, cont.

a **Log likelihood** - This is the log likelihood of the model. The value -1871.0093 has no meaning in and of itself; rather, this number can be used to help compare nested models.

b **Number of obs** - This is the number of observations that were used in the analysis. This number may be smaller than the total number of observations in your data set if you have missing values for any of the variables used in the logistic regression.

c **LR chi2(2)** - This is the likelihood ratio (LR) chi-square test. The number in the parenthesis indicates the number of degrees of freedom. In this model, there are 2 predictors, so there are 2 degrees of freedom.

d **Prob > chi2** - This is the probability of obtaining the chi-square statistic given that the null hypothesis is true. In other words, this is the probability of obtaining this chi-square statistic (16.31) if there is in fact no effect of the independent variables, taken together, on the dependent variable. In this case, the model is statistically significant because the p-value is 0.0003

e **Pseudo R2** - This is the pseudo R-squared. Logistic regression does not have an equivalent to the R-squared that is found in linear regression; however, many people have tried to come up with one. There are a wide variety of pseudo-R-square statistics. Because this statistic does not mean what R-square means in linear regression (the proportion of variance explained by the predictors), interpret this statistic with great caution.

Logistic regression annotated output, cont.

f heartatk - This is the outcome in our logistic regression. The variables listed below it are the predictors (exposure and covariates)

g Odds Ratio – These are the odds ratios for the logistic regression equation for predicting the outcome variable (headache) from the exposure (acupuncture) and covariate (baseline health score). They are converted from the log-odds units used for the logistic regression equation.

log(p/1-p) = b0 + b1*highbp+ b2*sex

g highbp– Study participants with high blood pressure had 1.7 times higher odds of heartatk compared to study partipants with lower blood pressure, controlling for sex

h sex– Females have .48 times lower odds of heat attack compared to males, controlling for blood pressure.

i constant - This is the expected value of the log-odds of **heart attack** when all of the predictor variables equal zero.

k Std. Err. - These are the standard errors associated with the ORs. The standard error is used for testing whether the parameter is significantly different from 0; by dividing the parameter estimate by the standard error you obtain a z-value (see the column with z-values and p-values). The standard errors can also be used to form a confidence interval for the parameter, as shown in the last two columns of this table.

l z and **P>|z|** - These columns provide the z-value and 2-tailed p-value used in testing the null hypothesis that the OR is 0.

m [95% Conf. Interval] - This shows a 95% confidence interval for the OR.