Statistics One

Lecture 8

Multiple Regression

Three segments

- Intro to multiple regression
- Matrix algebra
- Estimation of coefficients

Lecture 8
Segment 1

Intro to Multiple Regression

Multiple Regression

- Important concepts/topics
  - Multiple regression equation
  - Interpretation of regression coefficients
  - Standard vs. sequential regression

Simple vs. multiple regression

- Simple regression
  - Just one predictor (X)
- Multiple regression
  - Multiple predictors (X1, X2, X3, …)

Multiple regression

- Multiple regression equation
  - Just add more predictors (multiple Xs)

$$\hat{Y} = B_0 + B_1X_1 + B_2X_2 + B_3X_3 + \ldots + B_kX_k$$

$$\hat{Y} = B_0 + \Sigma(B_kX_k)$$

Multiple regression

- Multiple regression equation

  $\hat{Y}$ = predicted value on the outcome variable Y
  B_0 = predicted value on Y when all X = 0
  X_k = predictor variables
  B_k = unstandardized regression coefficients
  $Y - \hat{Y}$ = residual (prediction error)
  k = the number of predictor variables

Model R and R^2

- R = multiple correlation coefficient
  - $R = r_{\acute{Y}Y}$
  - The correlation between the predicted scores and the observed scores
- R^2
  - The percentage of variance in Y explained by the model

Multiple regression: Example

- Outcome measure (Y)
  - Faculty salary (Y)
- Predictors (X1, X2, X3)
  - Time since PhD (X1)
  - # of publications (X2)
  - Gender (X3)

Descriptive Statistics

	N	Mean	Std. Deviation	Skew	Kurtosis
SALARY	150	64,115.17	17,110.15	.25	-.55
TIME	150	8.09	5.24	.49	-.34
PUBS	150	15.49	7.51	.37	.17

Multiple regression: Example

- Gender
  - Male = 0
  - Female = 1

Multiple regression: Example

- $\hat{Y} = 46{,}911 + 1{,}382(\text{TIME}) + 502(\text{PUBS}) + -3{,}484(\text{G})$

Coefficients

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	p
(Constant)	46,910.49	3,401.423		13.791	.000
TIME	1382.07	235.980	.423	5.857	.000
PUBS	501.73	164.480	.220	3.050	.003
GENDER	-3,483.65	2,438.766	-.102	-1.428	.155

Model Summary

Model	R	R Square
1	.513	.263

Types of multiple regression

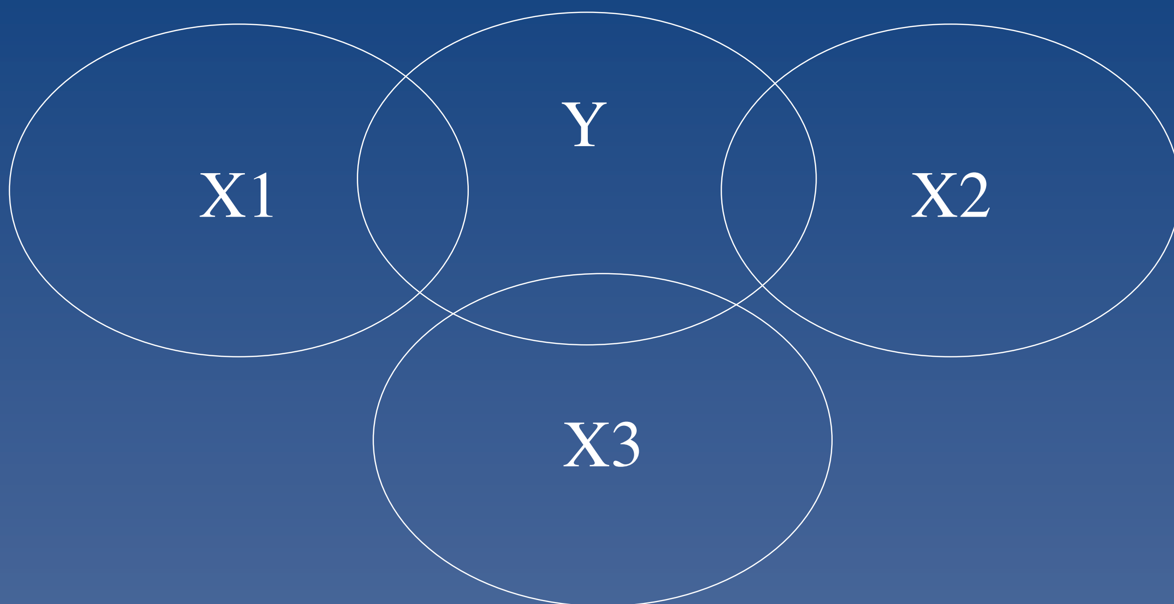
- Standard
- Sequential (aka hierarchical)

  – The difference between these approaches is how they handle the correlations among predictor variables

Types of multiple regression

- If X1, X2, and X3 are not correlated then type of regression analysis doesn't matter
  - See Venn diagram on next slide

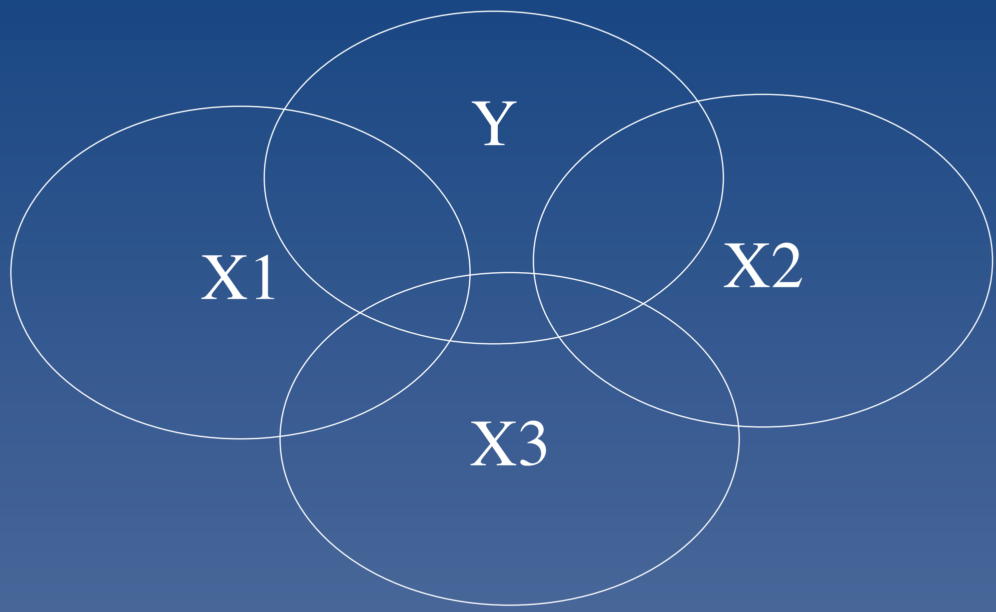
Orthogonal predictors



Types of multiple regression

- If predictors are correlated then different methods will return different results
  - See Venn diagram on next slide

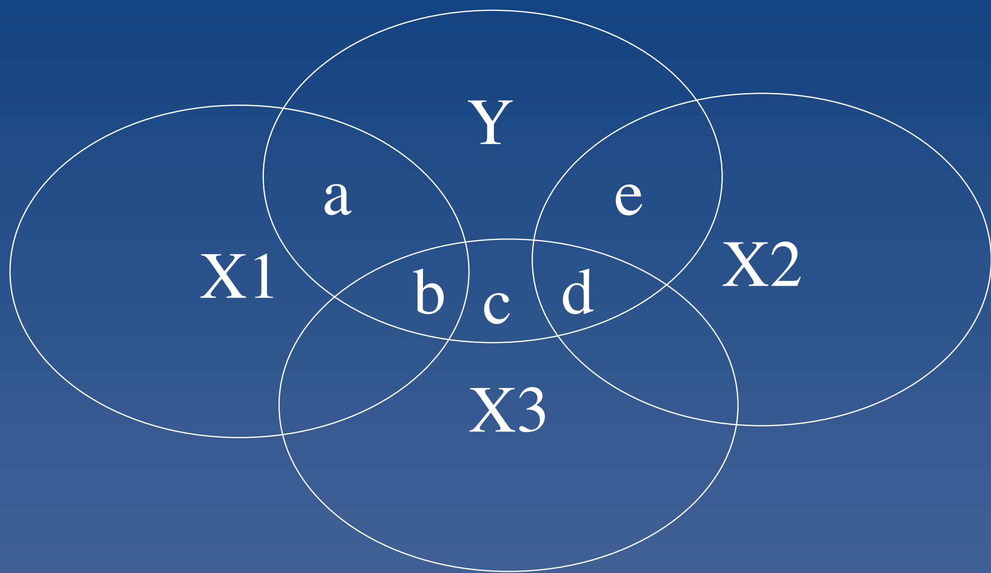
Correlated predictors



Standard

- All predictors are entered into the regression equation at the same time
- Each predictor is evaluated in terms of what it adds to the prediction of Y that is different from the predictability offered by the others
- Overlapping areas are assigned to R^2 but not to any individual B

Standard



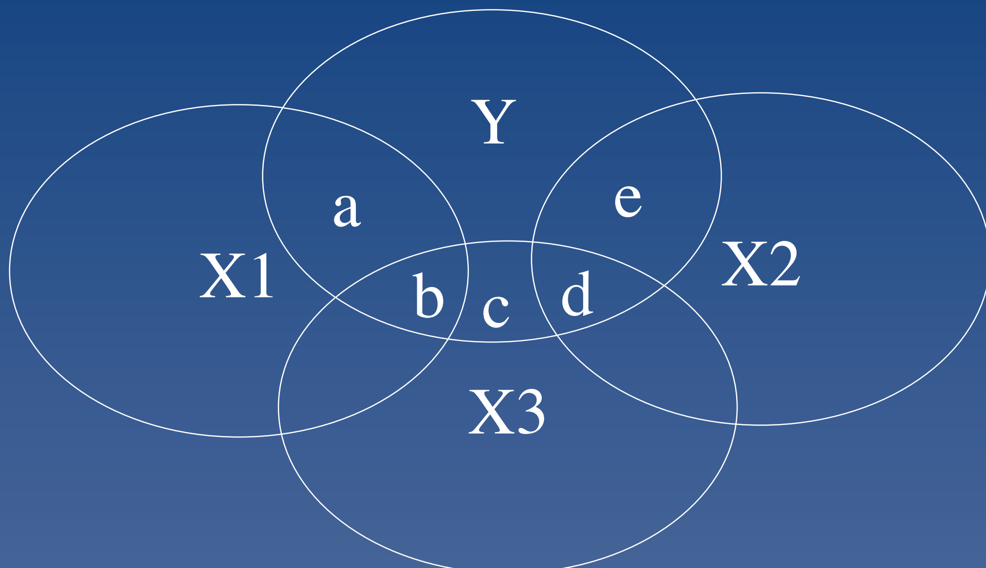
X1: a
X2: e
X3: c

Model R^2:
a+b+c+d+e

Sequential

- Predictors are entered into the regression equation in ordered steps; the order is specified by the researcher
- Each predictor is assessed in terms of what it adds to the equation *at its point of entry*
- Often useful to assess the change in R^2 from one step to another

Step 1: X1; Step 2: X2 + X3



X1: a + b
X2: e
X3: c

Step 1 Model R^2:
a+b

Step 2 Model R^2:
a+b+c+d+e

Multiple Regression

- Important concepts/topics
  - Multiple regression equation
  - Interpretation of regression coefficients
  - Standard vs. sequential regression