# Multivariate Multiple Linear Regression

Dr. Daniel B. Rowe
Professor of Computational Statistics
Department of Mathematical and Statistical Sciences
Marquette University

# Outline

**Univariate Simple Sampling**

**Univariate Simple Linear Regression**

**Univariate Multiple Linear Regression**

**Multivariate Multiple Linear Regression**

**Example**

# Univariate Simple Sampling

**Slides Reordered To See Symmetry**

Statistical model: Observe $1 \times 1$ scalar observations $0$ regressors

observation $i$

$$\underset{1\times 1}{y_i} = \underset{1\times 1}{\beta_0} + \underset{1\times 1}{\varepsilon_i} \qquad i = 1,\ldots,n \qquad E(\varepsilon_i) = 0 \qquad \mathrm{var}(\varepsilon_i) = \sigma^2$$

$\mu$

equivalently

observation $i$

$$\underset{1\times 1}{y_i} = \underset{1\times 1}{(1)}\underset{1\times 1}{\left(\beta_0\right)} + \underset{1\times 1}{\varepsilon_i}$$

or equivalently

# Univariate Simple Sampling

$$y_i = \beta_0 + \varepsilon_i$$

$$i = 1, \ldots, n$$

Statistical model: Observe $1 \times 1$ scalar observations 0 regressors

$$\underset{n\times 1}{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}} = \underset{n\times 1}{\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}} \underset{1\times 1}{\begin{pmatrix} \beta_0 \end{pmatrix}} + \underset{n\times 1}{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}}$$

$$i = 1, \ldots, n$$

$$E(\varepsilon_i) = 0$$

more compactly

$$\underset{n\times 1}{y} = \underset{n\times 1}{X} \underset{1\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

$$\text{var}(\varepsilon_i) = \sigma^2$$

4

# Univariate Simple Sampling

$$y_i = \beta_0 + \varepsilon_i$$

$$i = 1,\ldots,n$$

Statistical model: Observe $1 \times 1$ scalar observations 0 regressors

$$\underset{n\times 1}{y} = \underset{n\times 1}{X} \underset{1\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

$\mu$

with coefficient (mean) estimated as

$$\underset{1\times 1}{\hat{\beta}} = (\underset{1\times n}{X'} \underset{n\times 1}{X})^{-1} \underset{1\times n}{X'} \underset{n\times 1}{y} \qquad n \geq 1,$$

general rule $n > 10$

$\hat{\mu} = \bar{x}$

$$E(\varepsilon) = 0$$

from score function

$$\underset{1\times 1}{Q} = (y - X\beta)'(y - X\beta)$$

Exponent in multivariate normal PDF

$$\mathrm{cov}(\varepsilon) = \sigma^2 I_n$$

$$f(\underset{n\times 1}{y} \mid \beta, \sigma^2) = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)}$$

# Univariate Simple Sampling

$$\underset{n\times 1}{y} = \underset{n\times 1}{X} \underset{1\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

$\mu$

Statistical model: Observe $1 \times 1$ scalar observations 0 regressors

$$\underset{1\times 1}{\hat{\beta}} = (\underset{1\times n}{X'} \underset{n\times 1}{X})^{-1} \underset{1\times n}{X'} \underset{n\times 1}{y} \qquad n \geq 1, \qquad \text{general rule } n>10$$

$\hat{\mu} = \bar{x}$

$$\underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n-1}(y - X\hat{\beta})'(y - X\hat{\beta})$$

leads to

$$E(\underset{1\times 1}{\hat{\beta}}) = \beta \qquad \operatorname{cov}(\underset{1\times 1}{\hat{\beta}}) = \sigma^2 \left(X'X\right)^{-1}$$

$$E(\varepsilon) = 0$$

$$\operatorname{cov}(\varepsilon) = \sigma^2 I_n$$

and if $\varepsilon$ is normally distributed,

$\hat{\beta}$ is normally distributed and $(n-1)\hat{\sigma}^2$ is gamma!

$$(n-1)\hat{\sigma}^2 / \sigma^2 \sim \chi^2$$

# Univariate Simple Linear Regression

Statistical model: Observe $1 \times 1$ scalar observations 1 regressor

observation $i$

$$\underset{1\times 1}{y_i} = \underset{1\times 1}{\beta_0} + \underset{1\times 1}{\beta_1}\underset{1\times 1}{x_i} + \underset{1\times 1}{\varepsilon_i} \qquad i=1,\ldots,n \qquad E(\varepsilon_i)=0 \qquad \text{var}(\varepsilon_i)=\sigma^2$$

equivalently

observation $i$

$$\underset{1\times 1}{y_i} = \underset{1\times 2}{(1, x_i)} \underset{2\times 1}{\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}} + \underset{1\times 1}{\varepsilon_i}$$

or equivalently

# Univariate Simple Linear Regression

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

$$i = 1, \ldots, n$$

Statistical model: Observe $1 \times 1$ scalar observations 1 regressor

$$\underset{n\times 1}{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}} = \underset{n\times 2}{\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}} \underset{2\times 1}{\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}} + \underset{n\times 1}{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}}$$

$$i = 1, \ldots, n$$

$$E(\varepsilon_i) = 0$$

more compactly

$$\underset{n\times 1}{y} = \underset{n\times 2}{X} \; \underset{2\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

$$\operatorname{var}(\varepsilon_i) = \sigma^2$$

# Univariate Simple Linear Regression

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

$$i = 1, \ldots, n$$

Statistical model: Observe $1 \times 1$ scalar observations 1 regressor

$$\underset{n\times 1}{y} = \underset{n\times 2}{X} \; \underset{2\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

with coefficient (mean) estimated as

$$\underset{2\times 1}{\hat{\beta}} = (\underset{2\times n}{X'} \underset{n\times 2}{X})^{-1} \underset{2\times n}{X'} \underset{n\times 1}{y} \qquad n \geq 2, \qquad \text{general rule } n > 20$$

$$E(\varepsilon) = 0$$

from score function

$$\underset{1\times 1}{Q} = (y - X\beta)'(y - X\beta)$$

$$\operatorname{cov}(\varepsilon) = \sigma^2 I_n$$

Exponent in multivariate normal PDF

$$f(\underset{n\times 1}{y} \mid \beta, \sigma^2) = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)}$$

# Univariate Simple Linear Regression

$$\underset{n\times 1}{y} = \underset{n\times 2}{X}\ \underset{2\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

Statistical model: Observe $1\times 1$ scalar observations 1 regressor

$$\underset{2\times 1}{\hat{\beta}} = (\underset{2\times n}{X'}\underset{n\times 2}{X})^{-1}\underset{2\times n}{X'}\ \underset{n\times 1}{y} \qquad n\geq 2, \qquad \text{general rule } n>20$$

$$\underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n-2}(y - X\hat{\beta})'(y - X\hat{\beta})$$

leads to

$$E(\varepsilon)=0$$

$$\underset{2\times 1}{E(\hat{\beta})} = \beta \qquad \underset{2\times 2}{\mathrm{cov}(\hat{\beta})} = \sigma^2\left(X'X\right)^{-1}$$

$$\mathrm{cov}(\varepsilon) = \sigma^2 I_n$$

and if $\varepsilon$ is normally distributed,

$\hat{\beta}$ is normally distributed and $(n-2)\hat{\sigma}^2$ is gamma!

$$(n-2)\hat{\sigma}^2/\sigma^2 \sim \chi^2$$

# Univariate Multiple Linear Regression

Statistical model: Observe $1 \times 1$ scalar observations $q$ regressors

observation $i$

$$\underset{1\times 1}{y_i} = \underset{1\times 1}{\beta_0} + \underset{1\times 1}{\beta_1}\underset{1\times 1}{x_{i1}} + \ldots + \underset{1\times 1}{\beta_q}\underset{1\times 1}{x_{iq}} + \underset{1\times 1}{\varepsilon_i} \qquad i = 1,\ldots,n \qquad E(\varepsilon_i) = 0 \qquad \mathrm{var}(\varepsilon_i) = \sigma^2$$

equivalently

observation $i$

$$\underset{1\times 1}{y_i} = \underset{1\times(q+1)}{(1, x_{i1}, \ldots, x_{iq})} \underset{(q+1)\times 1}{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_q \end{pmatrix}} + \underset{1\times 1}{\varepsilon_i}$$

or equivalently

# Univariate Multiple Linear Regression

$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_q x_{iq} + \varepsilon_i$$
$$i = 1, \ldots, n$$

Statistical model: Observe $1 \times 1$ scalar observations $q$ regressors

$$\underset{n \times 1}{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}} = \underset{n \times (q+1)}{\begin{pmatrix} 1 & x_{11} & \cdots & x_{1q} \\ 1 & x_{21} & & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nq} \end{pmatrix}} \underset{(q+1) \times 1}{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_q \end{pmatrix}} + \underset{n \times 1}{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}}$$

$$i = 1, \ldots, n$$

$$E(\varepsilon_i) = 0$$

more compactly

$$\underset{n \times 1}{y} = \underset{n \times (q+1)}{X} \underset{(q+1) \times 1}{\beta} + \underset{n \times 1}{\varepsilon}$$

$$\mathrm{var}(\varepsilon_i) = \sigma^2$$

# Univariate Multiple Linear Regression

$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_q x_{iq} + \varepsilon_i$$
$$i = 1, \ldots, n$$

Statistical model: Observe $1 \times 1$ scalar observations $q$ regressors

$$\underset{n\times 1}{y} = \underset{n\times(q+1)}{X}\underset{(q+1)\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

with coefficient (mean) estimated as

$$\underset{(q+1)\times 1}{\hat{\beta}} = (\underset{(q+1)\times n}{X'}\underset{n\times(q+1)}{X})^{-1}\underset{(q+1)\times n}{X'}\underset{n\times 1}{y}$$

$n \geq (q+1)$, general rule $n > 10(q+1)$

$$E(\varepsilon) = 0$$

from score function

$$\underset{1\times 1}{Q} = (y - X\beta)'(y - X\beta)$$

$$\mathrm{cov}(\varepsilon) = \sigma^2 I_n$$

Exponent in multivariate normal PDF

$$f(\underset{n\times 1}{y} \mid \beta, \sigma^2) = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)}$$

# Univariate Multiple Linear Regression

$$\underset{n\times 1}{y} = \underset{n\times(q+1)}{X}\ \underset{(q+1)\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

Statistical model: Observe $1\times 1$ scalar observations $q$ regressors

$$\underset{(q+1)\times 1}{\hat{\beta}} = (\underset{(q+1)\times n}{X'}\underset{n\times(q+1)}{X})^{-1}\underset{(q+1)\times n}{X'}\ \underset{n\times 1}{y} \qquad n\geq(q+1), \text{ general rule } n>10(q+1)$$

$$\underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n-q-1}(y-X\hat{\beta})'(y-X\hat{\beta})$$

leads to

$$\underset{(q+1)\times 1}{E(\hat{\beta})} = \beta \qquad \underset{(q+1)\times(q+1)}{\mathrm{cov}(\hat{\beta})} = \sigma^2\left(X'X\right)^{-1}$$

$$E(\varepsilon) = 0$$

$$\mathrm{cov}(\varepsilon) = \sigma^2 I_n$$

and if $\varepsilon$ is normally distributed,

$\hat{\beta}$ is normally distributed and $(n-q-1)\hat{\sigma}^2$ is gamma!

$$(n-q-1)\hat{\sigma}^2/\sigma^2 \sim \chi^2$$

# Multivariate Multiple Linear Regression

Statistical model: Observe $1\times p$ vector observations $q$ regressors

$$\underset{1\times p}{\overset{\text{observation } i}{(y_{i1},..,y_{ip})}} = \underset{1\times p}{\begin{pmatrix} \beta_{01} & + & \beta_{11}x_{i1} & + & \cdots & + & \beta_{q1}x_{iq} \\ \beta_{02} & + & \beta_{12}x_{i1} & + & \cdots & + & \beta_{q2}x_{iq} \\ \vdots & + & \vdots & + & \ddots & + & \vdots \\ \beta_{0p} & + & \beta_{1p}x_{i1} & + & \cdots & + & \beta_{qp}x_{iq} \end{pmatrix}'} + \underset{1\times p}{(\varepsilon_{i1},...,\varepsilon_{ip})} \qquad i=1,...,n \qquad E(\varepsilon_i)=0 \qquad \text{cov}(\varepsilon_i')=\Sigma$$

equivalently

$$\underset{1\times p}{\overset{\text{observation } i}{(y_{i1},..,y_{ip})}} = \underset{1\times(q+1)}{(1,x_{i1},...,x_{iq})} \underset{(q+1)\times p}{\begin{pmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0p} \\ \beta_{11} & \beta_{12} & & \beta_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{pmatrix}} + \underset{1\times p}{(\varepsilon_{i1},...,\varepsilon_{ip})}$$

$$\varepsilon_i' = \underset{1\times p}{\begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \vdots \\ \varepsilon_{ip} \end{pmatrix}}$$

or equivalently

# Multivariate Multiple Linear Regression

$$(y_{i1},...,y_{ip}) = (1, x_{i1},...,x_{iq})\begin{pmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{10} \\ \beta_{11} & \beta_{12} & & \beta_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{pmatrix} + (\varepsilon_{i1} + ... + \varepsilon_{ip}) \quad i = 1,...,n$$

Statistical model: Observe $1 \times p$ vector observations $q$ regressors

$$\underset{n \times p}{\begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1p} \\ y_{21} & y_{22} & & y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{np} \end{pmatrix}} = \underset{n \times (q+1)}{\begin{pmatrix} 1 & x_{11} & \cdots & x_{1q} \\ 1 & x_{21} & & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nq} \end{pmatrix}} \underset{(q+1) \times p}{\begin{pmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0p} \\ \beta_{11} & \beta_{12} & & \beta_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{pmatrix}}$$

$$i = 1,...,n$$

$$+ \underset{n \times p}{\begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1p} \\ \varepsilon_{21} & \varepsilon_{22} & & \varepsilon_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} & \cdots & \varepsilon_{np} \end{pmatrix}}$$

$$E(\varepsilon_i) = 0$$

$$\text{cov}(\varepsilon_i') = \Sigma$$

more compactly

$$\underset{n \times p}{Y} = \underset{n \times (q+1)}{X} \underset{(q+1) \times p}{B} + \underset{n \times p}{E}$$

$$\underset{(q+1) \times p}{B} = (\underbrace{\beta_1,...,\beta_p}_{\text{cols}}) = (\underbrace{b_1,...,b_{q+1}}_{\text{rows}})'$$

# Multivariate Multiple Linear Regression

$$(y_{i1},...,y_{ip})=(1,x_{i1},...,x_{iq})\begin{pmatrix}\beta_{01} & \beta_{02} & \cdots & \beta_{10}\\ \beta_{11} & \beta_{12} & & \beta_{1p}\\ \vdots & \vdots & \ddots & \vdots\\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp}\end{pmatrix}+(\varepsilon_{i1}+...+\varepsilon_{ip})\quad i=1,...,n$$

Statistical model: Observe $1\times p$ vector observations $q$ regressors

$$\underset{n\times p}{Y}=\underset{n\times(q+1)}{X}\ \underset{(q+1)\times p}{B}+\underset{n\times p}{E}$$

with coefficient (mean) estimated as

$$\underset{(q+1)\times p}{\hat{B}}=(\underset{(q+1)\times n}{X'}\underset{n\times(q+1)}{X})^{-1}\underset{(q+1)\times n}{X'}\ \underset{n\times p}{Y}$$

$n\geq(q+1)p$, general rule $n>10p(q+1)$

$$E(E)=0$$

from score function

$$\underset{1\times 1}{Q}=tr(Y-XB)'(Y-XB)$$

Exponent in matrix normal PDF

$$\mathrm{cov}(vec(E'))=I_n\otimes\Sigma$$

Each row of $E$ has $\mathrm{cov}(\varepsilon_i)=\Sigma$

$$f(\underset{n\times p}{Y}\,|\,B,\Sigma)=(2\pi)^{-np/2}\,|\Sigma|^{-n/2}\,e^{-\frac{1}{2}tr\Sigma^{-1}(Y-XB)'(Y-XB)}$$

$tr(\cdot)$ is trace, sum of diagonal elements

# Multivariate Multiple Linear Regression

$$\underset{n\times p}{Y} = \underset{n\times(q+1)}{X}\;\underset{(q+1)\times p}{\beta} + \underset{n\times p}{E}$$

Statistical model: Observe $1\times p$ vector observations $q$ regressors

$$\underset{(q+1)\times p}{\hat{B}} = (\underset{(q+1)\times n}{X'}\underset{n\times(q+1)}{X})^{-1}\underset{(q+1)\times n}{X'}\;\underset{n\times p}{Y}$$ $n\geq(q+1)p$, general rule $n>10p(q+1)$

$$\underset{p\times p}{\hat{\Sigma}} = \frac{1}{n-q-1}(Y-X\hat{B})'(Y-X\hat{B})$$

leads to

$$E(E)=0$$

$$\underset{(q+1)\times p}{E(\hat{B})} = B \qquad \underset{p(q+1)\times p(q+1)}{\mathrm{cov}(vec(\hat{B}))} = \left(X'X\right)^{-1}\otimes\Sigma$$

$$\mathrm{cov}(\underset{\text{cols}}{\hat{\beta}_j}) = \sigma_j^2\left(X'X\right)^{-1}$$

$$\mathrm{cov}(vec(E')) = I_n\otimes\Sigma$$

$$\mathrm{cov}(\underset{\text{rows}}{\hat{b}_i}) = w_{ii}\Sigma$$

and if $E$ is normally distributed,

$$W = \left(X'X\right)^{-1}$$

$\hat{B}$ is normally distributed and $(n-q-1)\hat{\Sigma}$ is Wishart!

$$(n-q-1)\hat{\Sigma}\sim Wishart$$

# Multivariate Multiple Linear Regression

## Example:

It is believed that the heights of sons ($y_s$) and daughters ($y_d$) are linearly dependent upon the heights of their fathers ($x_f$) and mothers ($x_m$).

$$(y_{is}, y_{id}) = (1, x_{if}, x_{im}) \begin{pmatrix} \beta_{0s} & \beta_{0d} \\ \beta_{1s} & \beta_{1d} \\ \beta_{2s} & \beta_{2d} \end{pmatrix} + (\varepsilon_{is}, \varepsilon_{id}) \qquad (\varepsilon_{is}, \varepsilon_{id})' \sim N(0, \Sigma) \qquad i = 1, \ldots, n$$

Assume that the true parameter values are:

$$\underset{(q+1)\times p}{B} = \begin{pmatrix} 14.1 & 10.8 \\ 0.41 & 0.39 \\ 0.43 & 0.43 \end{pmatrix} \qquad \underset{p \times p}{\Sigma} = \begin{pmatrix} 2.250 & 1.125 \\ 1.125 & 2.250 \end{pmatrix}$$

could also imagine grandparents heights in model as $x$'s

http://chance.amstat.org/2013/09/1-pagano/

# Multivariate Multiple Linear Regression

## Example:

Selected

$x_f$ ={66,67,68,69,70,71,72,73,74,75},

$x_m$={60,61,62,63,64,65,66,67,68,69}

combinations ($n$=100), multiplied each with $B$

then added correlated normal noise with $\Sigma$.

Repeated $L$=10,000 times.

$$B = \begin{pmatrix} 14.1 & 10.8 \\ 0.41 & 0.39 \\ 0.43 & 0.43 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 2.250 & 1.125 \\ 1.125 & 2.250 \end{pmatrix}$$

$$A = \begin{pmatrix} 1.50 & 0 \\ 0.755 & 1.30 \end{pmatrix}$$

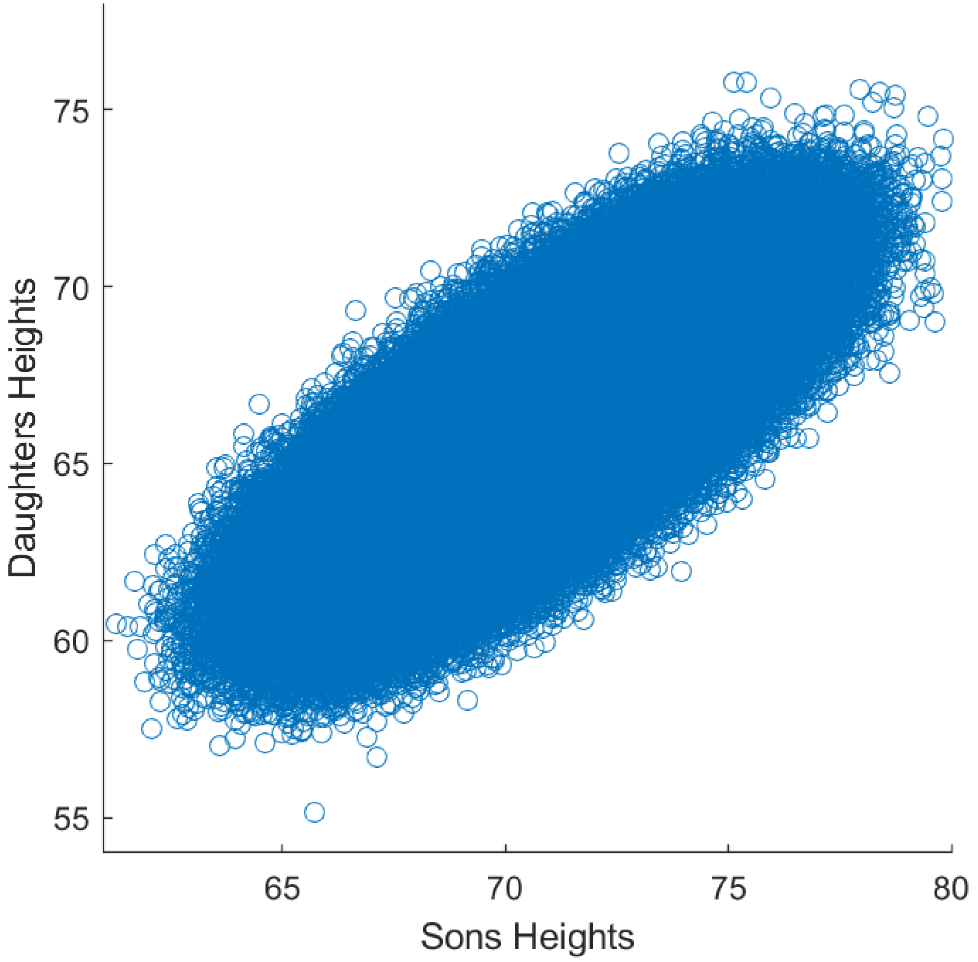
# Multivariate Multiple Linear Regression

## Example:

Estimates and plots from all $10^6$.

$$\hat{B} = \begin{pmatrix} 14.0511 & 10.8487 \\ 0.4108 & 0.3899 \\ 0.4299 & 0.4294 \end{pmatrix}$$

$$\hat{\Sigma} = \begin{pmatrix} 2.2504 & 1.1207 \\ 1.1207 & 2.2465 \end{pmatrix}$$




$$B = \begin{pmatrix} 14.1 & 10.8 \\ 0.41 & 0.39 \\ 0.43 & 0.43 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 2.250 & 1.125 \\ 1.125 & 2.250 \end{pmatrix}$$

$$A = \begin{pmatrix} 1.50 & 0 \\ 0.755 & 1.30 \end{pmatrix}$$

# Multivariate Multiple Linear Regression

```
% Multivariate model for
% Y1=son height, Y2=daughter height
% X1 = father height, X2=mother height
clear all
close all
rng('default')

% define true parameter values
n=100;
B=[[14.1;.41;.43],[10.8;.39;.43]]
sigma1=1.5; sigma2=1.5; rho=.5;
Sigma=[sigma1^2,sigma1*sigma2*rho;...
       sigma1*sigma2*rho,sigma2^2]
A=chol(Sigma)'

% generate simulated data
hf=(66:66+9);,hm=(60:60+9);
[Xf,Xm]=meshgrid(hf,hm);
xf=reshape(Xf,n,1);,xm=reshape(Xm,n,1);
X=[ones(n,1),xf,xm]
```

```
replicate=10000;
X=repmat(X,replicate,1);
n=size(X,1);

E=(A*randn(2,n))';
Y=X*B+E;

% Estimate parameters
Bhat=inv(X'*X)*X'*Y;
SigmaHat=(Y-X*Bhat)'*(Y-X*Bhat)/(n-3);

[Bhat,B]
[SigmaHat,Sigma]

figure;
scatter(Y(:,1),Y(:,2))
axis square
xlabel('Sons Heights')
ylabel('Daughters Heights')
xlim([61,80]),ylim([54,78])
```

# Univariate Simple Sampling

**Slides Reordered To See Symmetry**

Statistical model: Observe $1 \times 1$ scalar observations $0$ regressors

observation $i$

$$\underset{1\times 1}{y_i} = \underset{1\times 1}{\beta_0} + \underset{1\times 1}{\varepsilon_i} \qquad i = 1,\ldots,n \qquad E(\varepsilon_i) = 0 \qquad \text{var}(\varepsilon_i) = \sigma^2$$

$\mu$

equivalently

observation $i$

$$\underset{1\times 1}{y_i} = \underset{1\times 1}{(1)} \underset{1\times 1}{\left(\beta_0\right)} + \underset{1\times 1}{\varepsilon_i}$$

or equivalently

# Univariate Simple Linear Regression

Statistical model: Observe $1 \times 1$ scalar observations 1 regressor

observation $i$

$$\underset{1\times 1}{y_i} = \underset{1\times 1}{\beta_0} + \underset{1\times 1}{\beta_1}\,\underset{1\times 1}{x_i} + \underset{1\times 1}{\varepsilon_i} \qquad i = 1,\ldots,n \qquad E(\varepsilon_i) = 0 \qquad \mathrm{var}(\varepsilon_i) = \sigma^2$$

equivalently

observation $i$

$$\underset{1\times 1}{y_i} = \underset{1\times 2}{(1, x_i)} \underset{2\times 1}{\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}} + \underset{1\times 1}{\varepsilon_i}$$

or equivalently

# Univariate Multiple Linear Regression

Statistical model: Observe $1 \times 1$ scalar observations $q$ regressors

observation $i$

$$\underset{1\times 1}{y_i} = \underset{1\times 1}{\beta_0} + \underset{1\times 1}{\beta_1}\underset{1\times 1}{x_{i1}} + \ldots + \underset{1\times 1}{\beta_q}\underset{1\times 1}{x_{iq}} + \underset{1\times 1}{\varepsilon_i} \qquad i = 1,\ldots,n \qquad E(\varepsilon_i) = 0 \qquad \mathrm{var}(\varepsilon_i) = \sigma^2$$

equivalently

observation $i$

$$\underset{1\times q}{y_i} = \underset{1\times(q+1)}{(1, x_{i1}, \ldots, x_{iq})} \underset{(q+1)\times 1}{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_q \end{pmatrix}} + \underset{1\times 1}{\varepsilon_i}$$

or equivalently

# Multivariate Multiple Linear Regression

Statistical model: Observe $1\times p$ vector observations $q$ regressors

$$\underset{1\times p}{\overset{\text{observation } i}{(y_{i1},..,y_{ip})}} = \underset{1\times p}{\begin{pmatrix} \beta_{01} & + & \beta_{11}x_{i1} & + & \cdots & + & \beta_{q1}x_{iq} \\ \beta_{02} & + & \beta_{12}x_{i1} & + & \cdots & + & \beta_{q2}x_{iq} \\ \vdots & + & \vdots & + & \ddots & + & \vdots \\ \beta_{0p} & + & \beta_{1p}x_{i1} & + & \cdots & + & \beta_{qp}x_{iq} \end{pmatrix}'} + \underset{1\times p}{(\varepsilon_{i1},...,\varepsilon_{ip})} \qquad i=1,...,n \qquad E(\varepsilon_i)=0 \qquad \text{cov}(\varepsilon_i')=\Sigma$$

equivalently

$$\underset{1\times p}{\overset{\text{observation } i}{(y_{i1},..,y_{ip})}} = \underset{1\times(q+1)}{(1,x_{i1},...,x_{iq})} \underset{(q+1)\times p}{\begin{pmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0p} \\ \beta_{11} & \beta_{12} & & \beta_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{pmatrix}} + \underset{1\times p}{(\varepsilon_{i1},...,\varepsilon_{ip})}$$

$$\varepsilon_i' = \underset{1\times p}{\begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \vdots \\ \varepsilon_{ip} \end{pmatrix}}$$

or equivalently

# Univariate Simple Sampling

$$y_i = \beta_0 + \varepsilon_i$$

$$i = 1, \ldots, n$$

Statistical model: Observe $1 \times 1$ scalar observations 0 regressors

$$\underset{n \times 1}{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}} = \underset{n \times 1}{\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}} \underset{1 \times 1}{\left( \beta_0 \right)} + \underset{n \times 1}{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}}$$

$$i = 1, \ldots, n$$

$$E(\varepsilon_i) = 0$$

more compactly

$$\underset{n \times 1}{y} = \underset{n \times 1}{X} \underset{1 \times 1}{\beta} + \underset{n \times 1}{\varepsilon}$$

$$\text{var}(\varepsilon_i) = \sigma^2$$

27

# Univariate Simple Linear Regression

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

$$i = 1,\ldots,n$$

Statistical model: Observe $1 \times 1$ scalar observations 1 regressor

$$\underset{n\times 1}{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}} = \underset{n\times 2}{\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}} \underset{2\times 1}{\begin{pmatrix} \beta_0 \\ \beta_0 \end{pmatrix}} + \underset{n\times 1}{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}}$$

$$i = 1,\ldots,n$$

$$E(\varepsilon_i) = 0$$

more compactly

$$\underset{n\times 1}{y} = \underset{n\times 2}{X} \; \underset{2\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

$$\mathrm{var}(\varepsilon_i) = \sigma^2$$

# Univariate Multiple Linear Regression

$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_q x_{iq} + \varepsilon_i$$
$$i = 1, \ldots, n$$

Statistical model: Observe $1\times 1$ scalar observations $q$ regressors

$$\underset{n\times 1}{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}} = \underset{n\times(q+1)}{\begin{pmatrix} 1 & x_{11} & \cdots & x_{1q} \\ 1 & x_{21} & & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nq} \end{pmatrix}} \underset{(q+1)\times 1}{\begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_q \end{pmatrix}} + \underset{n\times 1}{\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}}$$

$$i = 1, \ldots, n$$

$$E(\varepsilon_i) = 0$$

more compactly

$$\underset{n\times 1}{y} = \underset{n\times(q+1)}{X} \underset{(q+1)\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

$$\text{var}(\varepsilon_i) = \sigma^2$$

# Multivariate Multiple Linear Regression

$$(y_{i1},...,y_{ip}) = (1, x_{i1},...,x_{iq}) \begin{pmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{10} \\ \beta_{11} & \beta_{12} & & \beta_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{pmatrix} + (\varepsilon_{i1} + ... + \varepsilon_{ip}) \quad i = 1,...,n$$

Statistical model: Observe $1 \times p$ vector observations $q$ regressors

$$\underset{n\times p}{\begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1p} \\ y_{21} & y_{22} & & y_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{np} \end{pmatrix}} = \underset{n\times(q+1)}{\begin{pmatrix} 1 & x_{11} & \cdots & x_{1q} \\ 1 & x_{21} & & x_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nq} \end{pmatrix}} \underset{(q+1)\times p}{\begin{pmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0p} \\ \beta_{11} & \beta_{12} & & \beta_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{pmatrix}} + \underset{n\times p}{\begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1p} \\ \varepsilon_{21} & \varepsilon_{22} & & \varepsilon_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} & \cdots & \varepsilon_{np} \end{pmatrix}}$$

$$i = 1,...,n$$

$$E(\varepsilon_i) = 0$$

$$\mathrm{cov}(\varepsilon_i') = \Sigma$$

more compactly

$$\underset{n\times p}{Y} = \underset{n\times(q+1)}{X} \underset{(q+1)\times p}{B} + \underset{n\times p}{E}$$

$$\underset{(q+1)\times p}{B} = (\underbrace{\beta_1,...,\beta_p}_{\text{cols}}) = (\underbrace{b_1,...,b_{q+1}}_{\text{rows}})'$$

# Univariate Simple Sampling

$$y_i = \beta_0 + \varepsilon_i$$

$$i = 1,\ldots,n$$

Statistical model: Observe $1 \times 1$ scalar observations 0 regressors

$$\underset{n\times 1}{y} = \underset{n\times 1}{X}\ \underset{1\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

$\mu$

with coefficient (mean) estimated as

$$\underset{1\times 1}{\hat{\beta}} = (\underset{1\times n}{X'}\underset{n\times 1}{X})^{-1}\underset{1\times n}{X'}\ \underset{n\times 1}{y} \qquad n\geq 1,$$

general rule $n>10$

$\hat{\mu} = \bar{x}$

$$E(\varepsilon) = 0$$

from score function

$$\underset{1\times 1}{Q} = (y - X\beta)'(y - X\beta)$$

Exponent in multivariate normal PDF

$$\mathrm{cov}(\varepsilon) = \sigma^2 I_n$$

$$f(\underset{n\times 1}{y}|\beta,\sigma^2) = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}(y-X\beta)'(y-X\beta)}$$

# Univariate Simple Linear Regression

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

$$i = 1,\ldots,n$$

Statistical model: Observe $1 \times 1$ scalar observations 1 regressor

$$\underset{n\times 1}{y} = \underset{n\times 2}{X}\ \underset{2\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

with coefficient (mean) estimated as

$$\underset{2\times 1}{\hat{\beta}} = (\underset{2\times n}{X'}\underset{n\times 2}{X})^{-1}\underset{2\times n}{X'}\underset{n\times 1}{y}$$ $n \geq 2$, general rule $n>20$

$$E(\varepsilon) = 0$$

from score function

$$\underset{1\times 1}{Q} = (y - X\beta)'(y - X\beta)$$

Exponent in multivariate normal PDF

$$\mathrm{cov}(\varepsilon) = \sigma^2 I_n$$

$$f(\underset{n\times 1}{y} \mid \beta, \sigma^2) = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)}$$

# Univariate Multiple Linear Regression

$$y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_q x_{iq} + \varepsilon_i$$
$$i = 1, \ldots, n$$

Statistical model: Observe $1 \times 1$ scalar observations $q$ regressors

$$\underset{n \times 1}{y} = \underset{n \times (q+1)}{X} \underset{(q+1) \times 1}{\beta} + \underset{n \times 1}{\varepsilon}$$

with coefficient (mean) estimated as

$$\underset{(q+1) \times 1}{\hat{\beta}} = (\underset{(q+1) \times n}{X'} \underset{n \times (q+1)}{X})^{-1} \underset{(q+1) \times n}{X'} \underset{n \times 1}{y}$$

$n \geq (q+1)$, general rule $n > 10(q+1)$

$$E(\varepsilon) = 0$$

from score function

$$\underset{1 \times 1}{Q} = (y - X\beta)'(y - X\beta)$$

Exponent in multivariate normal PDF

$$\operatorname{cov}(\varepsilon) = \sigma^2 I_n$$

$$f(\underset{n \times 1}{y} \mid \beta, \sigma^2) = (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}(y - X\beta)'(y - X\beta)}$$

# Multivariate Multiple Linear Regression

$$(y_{i1},...,y_{ip}) = (1, x_{i1},...,x_{iq})\begin{pmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{10} \\ \beta_{11} & \beta_{12} & & \beta_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{q1} & \beta_{q2} & \cdots & \beta_{qp} \end{pmatrix} + (\varepsilon_{i1} + ... + \varepsilon_{ip})$$

$$i = 1,...,n$$

Statistical model: Observe $1 \times p$ vector observations $q$ regressors

$$\underset{n\times p}{Y} = \underset{n\times(q+1)}{X}\,\underset{(q+1)\times p}{B} + \underset{n\times p}{E}$$

with coefficient (mean) estimated as

$$\underset{(q+1)\times p}{\hat{B}} = (\underset{(q+1)\times n}{X'}\,\underset{n\times(q+1)}{X})^{-1}\underset{(q+1)\times n}{X'}\,\underset{n\times p}{Y}$$

$n \geq (q+1)p$, general rule $n > 10p(q+1)$

$$E(E) = 0$$

from score function

$$\underset{1\times 1}{Q} = tr(Y - XB)'(Y - XB)$$

Exponent in matrix normal PDF

$$\mathrm{cov}(vec(E')) = I_n \otimes \Sigma$$

Each row of $E$ has $\mathrm{cov}(\varepsilon_i) = \Sigma$

$$f(\underset{n\times p}{Y} \mid B, \Sigma) = (2\pi)^{-np/2} \, |\Sigma|^{-n/2} \, e^{-\frac{1}{2} tr \Sigma^{-1}(Y-XB)'(Y-XB)}$$

$tr(\cdot)$ is trace, sum of diagonal elements

# Univariate Simple Sampling

$$\underset{n\times 1}{y} = \underset{n\times 1}{X}\ \underset{1\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

$\mu$

Statistical model: Observe $1\times 1$ scalar observations 0 regressors

$$\underset{1\times 1}{\hat{\beta}} = (\underset{1\times n}{X'}\underset{n\times 1}{X})^{-1}\underset{1\times n}{X'}\ \underset{n\times 1}{y} \qquad n\geq 1,$$

general rule $n>10$

$\hat{\mu} = \bar{x}$

$$\underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n-1}(y - X\hat{\beta})'(y - X\hat{\beta})$$

leads to

$$E(\varepsilon) = 0$$

$$\underset{1\times 1}{E(\hat{\beta})} = \beta \qquad \underset{1\times 1}{\mathrm{cov}(\hat{\beta})} = \sigma^2\left(X'X\right)^{-1}$$

$$\mathrm{cov}(\varepsilon) = \sigma^2 I_n$$

and if $\varepsilon$ is normally distributed,

$\hat{\beta}$ is normally distributed and $(n-1)\hat{\sigma}^2$ is gamma!

$$(n-1)\hat{\sigma}^2/\sigma^2 \sim \chi^2$$

# Univariate Simple Linear Regression

$$\underset{n\times 1}{y} = \underset{n\times 2}{X}\;\underset{2\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

Statistical model: Observe $1\times 1$ scalar observations 1 regressor

$$\underset{2\times 1}{\hat{\beta}} = (\underset{2\times n}{X'}\underset{n\times 2}{X})^{-1}\underset{2\times n}{X'}\;\underset{n\times 1}{y} \qquad n\geq 2, \qquad \text{general rule } n>20$$

$$\underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n-2}(y - X\hat{\beta})'(y - X\hat{\beta})$$

leads to

$$E(\varepsilon) = 0$$

$$\underset{2\times 1}{E(\hat{\beta})} = \beta \qquad \underset{2\times 2}{\mathrm{cov}(\hat{\beta})} = \sigma^2\left(X'X\right)^{-1}$$

$$\mathrm{cov}(\varepsilon) = \sigma^2 I_n$$

and if $\varepsilon$ is normally distributed,

$\hat{\beta}$ is normally distributed and $(n-2)\hat{\sigma}^2$ is gamma!

$$(n-2)\hat{\sigma}^2/\sigma^2 \sim \chi^2$$

# Univariate Multiple Linear Regression

$$\underset{n\times 1}{y} = \underset{n\times(q+1)}{X}\underset{(q+1)\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$$

Statistical model: Observe $1\times 1$ scalar observations $q$ regressors

$$\underset{(q+1)\times 1}{\hat{\beta}} = (\underset{(q+1)\times n}{X'}\underset{n\times(q+1)}{X})^{-1}\underset{(q+1)\times n}{X'}\underset{n\times 1}{y} \qquad n\geq(q+1), \text{ general rule } n>10(q+1)$$

$$\underset{1\times 1}{\hat{\sigma}^2} = \frac{1}{n-q-1}(y-X\hat{\beta})'(y-X\hat{\beta})$$

$$E(\varepsilon)=0$$

leads to

$$\underset{(q+1)\times 1}{E(\hat{\beta})} = \beta \qquad \underset{(q+1)\times(q+1)}{\text{cov}(\hat{\beta})} = \sigma^2\left(X'X\right)^{-1}$$

$$\text{cov}(\varepsilon) = \sigma^2 I_n$$

and if $\varepsilon$ is normally distributed,

$\hat{\beta}$ is normally distributed and $(n-q-1)\hat{\sigma}^2$ is gamma! $\qquad (n-q-1)\hat{\sigma}^2/\sigma^2 \sim \chi^2$

# Multivariate Multiple Linear Regression

$$\underset{n\times p}{Y} = \underset{n\times(q+1)}{X}\underset{(q+1)\times p}{B} + \underset{n\times p}{E}$$

Statistical model: Observe $1\times p$ vector observations $q$ regressors

$$\underset{(q+1)\times p}{\hat{B}} = (\underset{(q+1)\times n}{X'}\underset{n\times(q+1)}{X})^{-1}\underset{(q+1)\times n}{X'}\underset{n\times p}{Y}$$

$n\geq(q+1)p$, general rule $n>10p(q+1)$

$$\underset{p\times p}{\hat{\Sigma}} = \frac{1}{n-q-1}(Y-X\hat{B})'(Y-X\hat{B})$$

leads to

$$E(E)=0$$

$$\underset{(q+1)\times p}{E(\hat{B})} = B \qquad \underset{p(q+1)\times p(q+1)}{\operatorname{cov}(vec(\hat{B}))} = \left(X'X\right)^{-1}\otimes\Sigma$$

$$\operatorname{cov}(\underset{\text{cols}}{\hat{\beta}_j}) = \sigma_j^2\left(X'X\right)^{-1}$$

$$\operatorname{cov}(\underset{\text{rows}}{\hat{b}_i}) = w_{ii}\Sigma$$

$$\operatorname{cov}(vec(E')) = I_n\otimes\Sigma$$

$$W = \left(X'X\right)^{-1}$$

and if $E$ is normally distributed,

$\hat{B}$ is normally distributed and $(n-q-1)\hat{\Sigma}$ is Wishart!

$$(n-q-1)\hat{\Sigma}\sim Wishart$$

# Discussion

## Questions?

# Homework 7

1. Multivariate multiple linear regression

Select father $x_{if}$ and $x_{im}$ heights, multiply by coefficients $B$, add correlated error with $\Sigma$, repeat $n$=100 times.

$$(y_{is}, y_{id}) = (1, x_{if}, x_{im}) \begin{pmatrix} \beta_{0s} & \beta_{0d} \\ \beta_{1s} & \beta_{1d} \\ \beta_{2s} & \beta_{2d} \end{pmatrix} + (\varepsilon_{is}, \varepsilon_{id}) \qquad i = 1, \ldots, n$$

form $Y$ and $X$, estimate $\hat{B}$ and $\hat{\Sigma}$.

Repeat a large number of times $L$ so that you have $\hat{B}^{(1)}, \ldots, \hat{B}^{(L)}$ and $\hat{\Sigma}^{(1)}, \ldots, \hat{\Sigma}^{(L)}$. Compute means, variances, make histograms, etc.