STA 302 / 1001

Note Title 10/27/2011

We're skipping chapter 4 for now.

Tests will be returned on Tuesday @ ~~11:00~~ ??? end of lecture → either here or in Cquest

Assignment 2 is ready for you.

Andriy in RW 107/109 to help with A2:

Tues Nov. 1 11 – 1

Thus Nov 8 10 – 12 (Fall Break)

Wed Nov 9 10 – 12

Interpreting log-transformed data

"log" = ln

If $\log Y = \beta_0 + \beta_1 x + e$

$Y = e^{\beta_0} e^{\beta_1 x} e^{e}$

Note that errors are multiplicative rather than additive

Increases in x of 1 unit is associated with multiplicative change in Y of factor of e^{β_1}

Example: Breakdown $\widehat{\log Y} = 19. - .51$ Voltage

So 1 kV increase in voltage changes estimated mean of Y factor of $e^{-.51} = .6$

So voltage goes from 27 kV $\to$ 28 kV, time to breakdown is estimated to be 60% of what it was.

What if we transformed x?

$$Y = \beta_0 + \beta_1 \log(x) + e$$

Interpretation is in terms of multiplicative changes in x

For each k-fold change in x, estimated change in mean of Y is $\beta_1 \log k$

$$\left.\begin{array}{l} E(Y_{original}) = \beta_0 + \beta_1 \log(x) \\ E(Y_{new}) = \beta_0 + \beta_1 \log(kx) \end{array}\right\} E(Y_{original}) - E(Y_{new}) = \beta_1 \log k$$

What to do if assumptions are violated:

① Abandon simple linear regression for something else
- usually more complicated

eg. • Mauna Loa CFCs – need time series
- Weighted least squares for unequal variance (Ch. 4)
- models that allow non-normal errors
- robust methods – reduce effect of outliers

② For <u>outliers / influential</u> points

- check for errors
- If because points are from tails of skewed distribution consider transformation
- robust regression to reduce influence of outliers
- Consider reporting results with and without outliers

- consider a different model (non-linear) that also captures the unusual points

3. **Transformations**
   1. for non-linearity
   2. non-constant variance.
   3. non-normality

Transform X – may create linear transformation from non-linear

- if X is very right skewed, do a log or $\sqrt{\ }$ or $1/x$ transformation

Transform Y – for 3 reasons above

If e's (Y's) <u>not normal</u>:

CLT says linear combinations of r.v.'s are normally distributed even if original r.v.'s aren't. Our estimators of β_0, β_1 are linear combinations of r.v.'s so tests and CIs for them are robust against non-normality, as long as not too skewed or extreme outliers

Prediction intervals are NOT robust against non-normality → interval for <u>one</u> r.v.

Relative Importance of Assumptions for Inference

In order from most important to least:

- Right form of model ($E(e)=0$)
- Independence
- Equal variance – regression is reasonably robust against non-equal var. if have similar number of observations for each x
- Normality (necessary for PIs)

← END OF CHAPTER 3 →

CHAPTER 5:

Simple Linear Regression in Matrix Terms

Random vector. Suppose $X_1, \dots, X_n$ is a set of r.v.'s

$Y_1, \dots, Y_n$ is another set of r.v.'s

The random vector, $\mathbf{X}$ (bold)

my notation: X, $\underset{\sim}{X}$

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}$$ is a column vector

$$E(\underset{\sim}{X}) = \begin{pmatrix} E(X_1) \\ \vdots \\ E(X_2) \end{pmatrix}$$

If a is a scalar constant, $a\underset{\sim}{X} = \begin{pmatrix} aX_1 \\ aX_2 \\ \vdots \\ aX_n \end{pmatrix}$

$$E(a\underset{\sim}{X}) = a\,E(\underset{\sim}{X})$$

If A is a matrix of constants:

$$\underset{q \times n}{A}\ \underset{n \times 1}{\underset{\sim}{X}} \qquad E(A\underset{\sim}{X}) = A\,E(\underset{\sim}{X})$$

$$E(\underset{\sim}{X} + \underset{\sim}{Y}) = E(\underset{\sim}{X}) + E(\underset{\sim}{Y})$$

$\underset{\sim}{X}'$ (Both Sheather and I use $'$ (prime) for transpose)

is the row vector $(X_1, \ldots, X_n)$

If $\underset{\sim}{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$

If $\underset{\sim}{a}'\underset{\sim}{X} = a_1 X_1 + \dots + a_n X_n$

$$E(\underset{\sim}{a}'\underset{\sim}{X}) = \underset{\sim}{a}' E(\underset{\sim}{X})$$

Variance-Covariance Matrix, Var Matrix, Covariance Matrix

$Cov(\underset{\sim}{X})$, $Var(\underset{\sim}{X})$

$$Var(\underset{\sim}{X}) = E\left[(\underset{\sim}{X} - E(\underset{\sim}{X}))(\underset{\sim}{X} - E(\underset{\sim}{X}))'\right]_{n \times n}$$

$$= E\begin{bmatrix} (X_1 - E(X_1))^2 & (X_1 - EX_1)(X_2 - EX_2) & \cdots \\ \vdots & (X_2 - EX_2)^2 & \cdot \quad \cdot \\ & & \ddots \\ (X_1 - EX_1)(X_n - EX_n) & \vdots & (X_n - EX_n)^2 \end{bmatrix}$$

The i^{th} diagonal element is $Var(X_i)$

The ij^{th} element is $Cov(X_i, X_j)$

Since $Cov(X_i, X_j) = Cov(X_j, X_i)$

this is a <u>symmetric</u> matrix

For a symmetric matrix A, $A' = A$

Example

$$\underset{\sim}{e} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

e_i's errors in a regression model

$$E(\underset{\sim}{e}) = \underset{\sim}{0}$$

$$Var(\underset{\sim}{e}) = \begin{bmatrix} \sigma^2 & 0 & 0 & \cdots \\ & \sigma^2 & & 0 \\ & & \ddots & \\ 0 & & & \sigma^2 \end{bmatrix}$$

$$= \sigma^2 \underset{\sim}{I}$$

where $\underset{\sim}{I}$ is $n \times n$ identity matrix

$$I = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

$\text{Var}(A\underset{\sim}{X})$ $\quad$ A matrix of scalars

$$= E\left[(AX - E(AX))(AX - E(AX))' \right]$$

$$= E\left[A(X - E(X))(X - E(X))' A' \right]$$

$$= A\, E\left[(X - EX)(X - EX)' \right] A'$$

$$= A\, \text{Var}(\underset{\sim}{X})\, A'$$

$(A-B)' = A' - B'$

$(AB)' = B'A'$

Simple Linear Regression in Matrix Terms

Define the following vectors and matrices:

$$\underset{\sim}{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}_{n \times 1}, \quad \underset{\sim}{e} = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}_{n \times 1}, \quad \underset{\sim}{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}_{2 \times 1}$$

$$\underset{\sim}{X} = \begin{pmatrix} 1 & X_1 \\ 1 & \vdots \\ \vdots & \vdots \\ 1 & X_n \end{pmatrix}_{n \times 2}$$

Model: $\underset{\sim}{Y} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{e}$

Gauss-Markov conditions: $E(\underset{\sim}{e}) = \underset{\sim}{0}$

$Var(\underset{\sim}{e}) = \sigma^2 \underset{\sim}{I}$

↳ 2 conditions in 1

$\underset{\sim}{e}$ has n-dimensional Normal distribution

$$E(Y \mid X) = X\beta$$

$$Var(Y \mid X) = \sigma^2 \underset{\sim}{I}$$

Matrix Differentiation

If $\underset{\sim}{\theta} = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_k \end{pmatrix}$ and $f(\underset{\sim}{\theta})$ is a scalar

$$\frac{\partial f(\underset{\sim}{\theta})}{\partial \underset{\sim}{\theta}} = \begin{pmatrix} \partial f(\theta)/\partial \theta_1 \\ \vdots \\ \partial f(\theta)/\partial \theta_k \end{pmatrix}$$

Lemma: Suppose $\underset{\sim}{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix}$ and $f(\underset{\sim}{\theta}) = \underset{\sim}{c}' \underset{\sim}{\theta}$

$= \sum c_i \theta_i$

$$\partial f(\underset{\sim}{\theta}) / \partial \underset{\sim}{\theta} = \underset{\sim}{c} \quad //$$

Lemma 2 Let A be a $k \times k$ symmetric matrix

$$f(\underset{\sim}{\theta}) = \underset{\sim}{\theta}' A \underset{\sim}{\theta}$$

$$\frac{\partial f(\underset{\sim}{\theta})}{\partial \underset{\sim}{\theta}} = 2 A \underset{\sim}{\theta} \quad (k \times 1)$$

Least Squares Estimation of Regression Coefficients

Find β_0, β_1 that minimize the sum of squares of residuals

$$RSS(\underset{\sim}{\beta}) = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

$$= (Y - X\beta)'(Y - X\beta)$$

- Do the matrix multiplication to verify this

$$= (Y' - \beta' X')(Y - X\beta)$$

$$= Y'Y - \beta' X' Y - \underbrace{Y' X \beta}_{(1\times n)(n\times 2)(2\times 1)} + \beta' X' X \beta$$

$$= Y'Y - 2\beta' X' Y + \beta' X' X \beta$$

$$\partial RSS / \partial \tilde{b} = 0 - 2 \underset{2\times n}{X'} \underset{n\times 1}{Y} + 2 \underset{2\times 2}{X'X} \underset{2\times 1}{\beta}$$

($X'X$ is symmetric)

$$(X'X)' = X'(X')' = X'X$$

Set equal to 0

$$\cancel{2} X'X \hat{\beta} = \cancel{2} X'Y$$

$$\hat{\beta} = (X'X)^{-1} X'Y$$

- as long as $X'X$ is invertible
  $\det(X'X) \neq 0$
- rows / columns of $X'X$ are linearly

independent

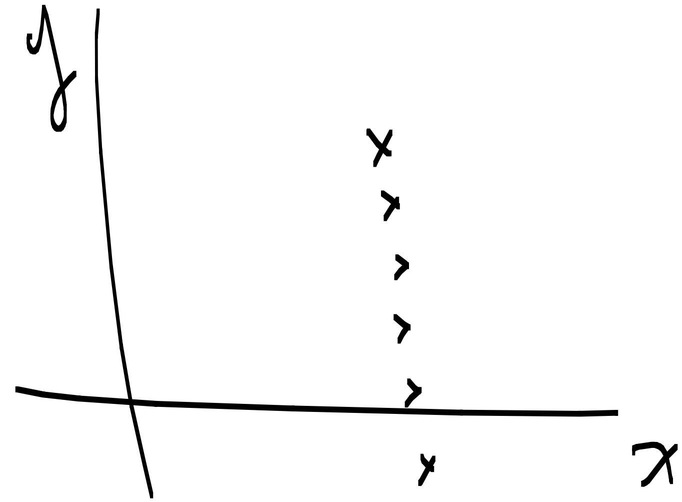
- rank of $X'X$ is 2

$$\text{rank}(AB) \leq \min(\text{rank } A, \text{rank } B)$$

$$X = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

If $\text{rank } X \neq 2$

$$x_1 = x_2 = \cdots = x_n$$



We'll assume this is not the case

Then $\operatorname{rank} X = 2$

$\operatorname{rank}(X'X) = 2$

What is $\underset{2\times 2}{X'X}$?

$$X'X = \begin{bmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & n\bar{x} \\ n\bar{x} & \sum x_i^2 \end{bmatrix}$$

symmetric ✓

What is $(X'X)^{-1}$?

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$