Generalized Linear, Non-linear and Mixed Models

Theory, Methods and Practice

December 6, 2015

- What is regression? What does regression do?

- What is regression? What does regression do?
- Regression ingredient:
  1. Response or dependent variable, usually denoted by Y; y
  2. Auxiliary or independent variable, usually denoted by X; x

- What is regression? What does regression do?
- Regression ingredient:
  1. Response or dependent variable, usually denoted by Y; y
  2. Auxiliary or independent variable, usually denoted by X; x
- Both Y and X may be univariable or multivariable.

- What is regression? What does regression do?
- Regression ingredient:
  1. Response or dependent variable, usually denoted by Y; y
  2. Auxiliary or independent variable, usually denoted by X; x
- Both Y and X may be univariable or multivariable.
- Dual Purpose:
  1. Develop relationship between Y and X;
  2. Predict future or unobserved Y using known or observed X.

- Deterministic world: $Y = f(x)$, f is completely known function.

  e.g. $A = \pi r^2$, $Y = A$, $x = r$, $f = \pi r^2$.

- Deterministic world: $Y = f(x)$, f is completely known function.

  e.g. $A = \pi r^2$, $Y = A$, $x = r$, $f = \pi r^2$.

- Indeterministic world or uncertain events

  $Y \approx (x)$

  or, $Y = f(x) + \varepsilon$.

- Deterministic world: $Y = f(x)$, f is completely known function.

  e.g. $A = \pi r^2$, $Y = A$, $x = r$, $f = \pi r^2$.

- Indeterministic world or uncertain events

  $Y \approx (x)$

  or, $Y = f(x) + \varepsilon$.

- Two components:

  f: Mathematical world;

  ε: Statistical world.

- However, f and ε could be related (How?).

- However, f and ε could be related (How?).
- Not necessarily statistically rather in terms of mathematical measurements

- However, f and ε could be related (How?).
- Not necessarily statistically rather in terms of mathematical measurements
- Linear Regression: $Y = \beta_0 + \beta_1 x + \varepsilon$.

- However, f and ε could be related (How?).
- Not necessarily statistically rather in terms of mathematical measurements
- Linear Regression: $Y = \beta_0 + \beta_1 x + \varepsilon$.
- What are the basic assumptions made?
  1. For every given x, Y is a random variable.
  2. The mean of Y is linearly related to x and in particular it is a straight line equation.

Other usual assumptions:

1. ε's are independent and identically distributed (iid).

Other usual assumptions:

1. ε's are independent and identically distributed (iid).
2. ε's are normally distributed: why this is required? In an indeterministic world measurement of uncertainty is the main thing and the uncertainty is measured by probability. The normality assumptions provide accurate calculation of uncertainty associated the Y and x relationship.

- Interpretation of β's: Important for many practical applications.

- Interpretation of β's: Important for many practical applications.
- How do you determine β? Why do they need to be fixed?

- Interpretation of β's: Important for many practical applications.
- How do you determine β? Why do they need to be fixed?
- For a given data set many straight lines seem plausible but we want one that uniquely determine the relationship. The uniqueness is defined by minimum squared error loss, in mathematical sense.

- However, they are uniquely determined by maximum likelihood estimation or some other methods (will be learned in this course) in statistical sense. Although they could be identical in many special cases.

- However, they are uniquely determined by maximum likelihood estimation or some other methods (will be learned in this course) in statistical sense. Although they could be identical in many special cases.
- Minimum Distance method

$$min_\beta \sum_{i=1}^{n}(y_i - \beta_0 - \beta_x x_i^2) \quad (1)$$

- ML method

$$max_\beta logL(\beta) = -\frac{n}{2}log(2\pi) - \frac{n}{2}log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_2 x_i)^2 \tag{2}$$

- ML method

$$max_{\beta} logL(\beta) = -\frac{n}{2}log(2\pi) - \frac{n}{2}log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_2 x_i)^2 \tag{2}$$

- You are required to know the statistical properties of the estimators of β

- Consider the data set extracted from a national study of 15 and 16 year old adolescents. The event of interest is ever having sexual intercourse.

Race	Gender	Intercourse	
		Yes	No
White	Male	43	134
	Female	26	149
Black	Male	29	23
	Female	22	36

-

$$\pi = Pr(intercourse)$$
$$= \frac{43+26+29+22}{(43+26+29+22)+(134+149+23+36)} = 0.260 \quad (3)$$

- $$\pi = Pr(intercourse) = \frac{43+26+29+22}{(43+26+29+22)+(134+149+23+36)} = 0.260 \tag{3}$$

- $$Pr(intercourse|white) = \frac{43+26}{43+26+134+149} = 0.196 \tag{4}$$

- $$Pr(intercourse|Black) = \frac{29 + 22}{29 + 22 + 23 + 36} = 0.463 \qquad (5)$$

- $$Pr(intercourse|Black) = \frac{29+22}{29+22+23+36} = 0.463 \tag{5}$$

- $$Pr(intercourse|Male) = \frac{43+29}{43+29+134+23} = 0.314 \tag{6}$$

- $$Pr(intercourse|Black) = \frac{29+22}{29+22+23+36} = 0.463 \qquad (5)$$

- $$Pr(intercourse|Male) = \frac{43+29}{43+29+134+23} = 0.314 \qquad (6)$$

- $$Pr(intercourse|Female) = \frac{26+22}{26+22+149+36} = 0.206 \qquad (7)$$

- $$Pr(intercourse|White\&Male) = \frac{43}{43+134} = 0.243 \qquad (8)$$

- $$Pr(intercourse|White\&Male) = \frac{43}{43+134} = 0.243 \tag{8}$$

- $$Pr(intercourse|White\&Female) = \frac{26}{26+149} = 0.148 \tag{9}$$

- $$Pr(intercourse|Black\&Male) = \frac{29}{29+23} = 0.558 \quad (10)$$

- $$Pr(intercourse|Black\&Male) = \frac{29}{29+23} = 0.558 \tag{10}$$

- $$Pr(intercourse|Black\&Female) = \frac{22}{22+36} = 0.379 \tag{11}$$

- How about having a formula: $\pi(x) = f(x)$?

  what x?

- How about having a formula: $\pi(x) = f(x)$?

  what x?

- Define conveniently

$$x_2 = \begin{cases} 1 & \text{if Male} \\ 0 & \text{if Female} \end{cases}$$

$$x_1 = \begin{cases} 1 & \text{if White} \\ 0 & \text{if Black} \end{cases}$$

- Let $f(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 = -0.455 - 0.1313x_1 + 0.648x_2$.

X		π(x)	
1	1	0.329	White Male
1	0	0.204	White Female
0	1	0.646	Black Male
0	0	0.488	Black Female

Issues:

- Structural defect-Probabilities fall between 0 and 1, where as linear functions take values over the entire real line. The model can be valid for a specific range of x, but not for all values of x.

Require a more general model formation.