Research Methods in Political Science I

9. Logistic Regression

Yuki Yanai

School of Law and Graduate School of Law

December 2, 2015

KOBE UNIVERSITY

Today's Menu

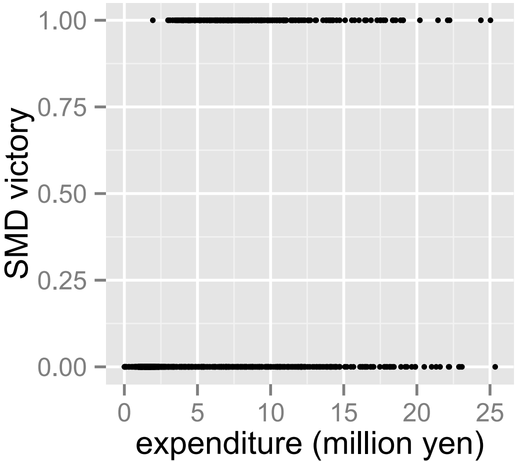
1. Logistic Regression
   - Introduction
   - Logistic Regression with One Predictor

2. Interpreting Logistic Regression
   - Estimated Coefficients
   - Statistical Inference

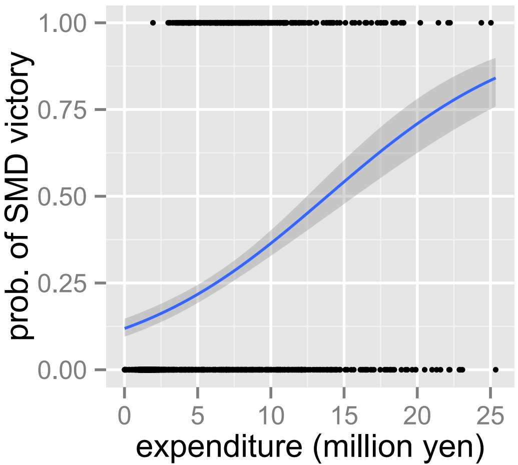
When We Use Logistic Regression

- Response is binary
- Response: $y_i = 0$ or 1 for all i
- E.g.
  - whether a person supports the current cabinet or not
  - whether a person turns out to vote or not
  - whether an armed conflict occurs or not
  - etc.

Visualize Binary Variable



Fitting the Logistic Curve



An Example Model of Logistic Regression

- Response: the result of SMD election (win or lose): $y_i =$0 (lost) or 1 (won)
- Predictor: electoral expenditure (million yen), `expm`
- Logistic curve of the model

$$\Pr(y_i = 1) = \text{logit}^{-1}(-2.00 + 0.14 \cdot \text{expm})$$

Logistic Regression Model

- Value of the response is either 0 or 1: linear models "$X\beta$ + error" don't fit well
- Instead, we model the probability of y being 1

$$\Pr(y_i = 1) = \text{logit}^{-1}(X_i\beta)$$

- Assumption: Given the probability of success p_i, each y_i is independently determined

$$y_i \sim \text{Bernoulli}(p_i)$$

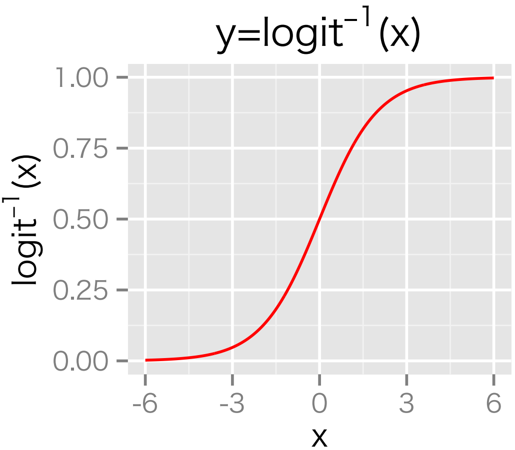
- We call $X\beta$ **linear predictor**（線形予測子）

Logistic [Inverse Logit] Function)

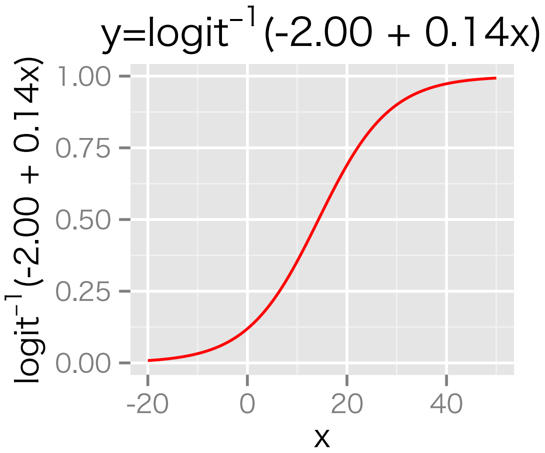
$$\operatorname{logit}^{-1}(x) = \frac{e^x}{1+e^x} = \frac{1}{1+\exp(-x)}$$

- Function that maps $x \in (-\infty, \infty)$ on (0, 1)
- Appropriate to treat probabilities
- Logistic function is the inverse of the logit function, hence we write $\operatorname{logit}^{-1}$

Logistic Curves (1)



Logistic Curves (2)



Logistic Function and Logit Function (1)

Logit function

- the inverse of logistic function
- maps a continuous variable $z \in (0,1)$ on $(-\infty, \infty)$

$$x_i = \mathrm{logit}(p_i) = \log\left(\frac{p_i}{1-p_i}\right)$$

Logistic Function and Logit Function

Two representations of the model

$$\Pr(y_i = 1) = p_i$$

1. Logistic

$$p_i = \text{logit}^{-1}(X_i\beta) = \frac{1}{1 + \exp(-X_i\beta)}$$

2. Logit

$$\text{logit}(p_i) = \log\left(\frac{p_i}{1 - p_i}\right) = X_i\beta$$

Both are important and frequently used

Logistic "Curves"

- Logistic curves are not lines: effects are not constant
- the amount of change in the response corresponding to 1-unit change of a predictor is not constant: **the effect size depends on the value of the predictor**
  - logit(0.5) = 0, logit(0.6) = 0.4: 0.4 unit change on logit scale is equivalent to 10 point change from 50% to 60% on probability scale
  - logit(0.07) = -2.6, logit(0.1) = -2.2: 0.4 unit change on logit scale is equivalent to 3 point change from 7% to 10% on probability scale
- the change in the response corresponding to a certain amount of change in the predictor: small when the predictor takes on values near minimum or maximum, and large when it takes on values around the mean

Evaluating the Data around the Mean (1)

$$\Pr(\text{win}) = \text{logit}^{-1}(-2.00 + 0.14 \cdot \text{expm})$$

- Logistic curve is not linear: the effect size depends on where of the data we evaluate
- Evaluate the probability of success at the mean
  - `mean(exam)` $= 6.12$
  - $\text{logit}^{-1}(-2.00 + 0.14 \cdot 6.12) = 0.24$

Evaluating the Data around the Mean (2)

$$\Pr(\text{win}) = \text{logit}^{-1}(-2.00 + 0.14 \cdot \text{expm})$$

- Evaluate the effect of the predictor on the success probability around the mean
  - `mean(exam)` $= 6.12$: Compare when expm = 6 and when expm = 7
  - $\text{logit}^{-1}(-2.00 + 0.14 \cdot 7) - \text{logit}^{-1}(-2.00 + 0.14 \cdot 6) = 0.026$
  - 1-unit increase of the expenditure *around its mean* leads to 2.6 point increase of the winning probability

Evaluating the Data around the Mean (3)

$$\Pr(\text{win}) = \text{logit}^{-1}(-2.00 + 0.14 \cdot \text{expm})$$

- Evaluate analytically,

$$\frac{d}{dx}\text{logit}^{-1}(-2+0.14x) = 0.14\frac{\exp(-2+0.14x)}{[1+\exp(-2+0.14x)]^2}$$

- Plug $\bar{x} = 6.12$ into x

$$\frac{d}{dx}\text{logit}^{-1}(-2+0.14\cdot 6.12) = 0.026$$

Divide-by-4 Rule

- the slope of the tangent of the logistic curve: max when $X\beta = \alpha + \beta x = 0$
- $\text{logit}^{-1}(0) = 0.5$
- the slope is

$$\frac{d}{dx}\text{logit}^{-1}(0) = \beta \frac{e^0}{(1+e^0)^2} = \frac{\beta}{4}$$

- Maximal effect of the predictor is the value of the coefficient divided by 4
- Winning probability: the coefficient of the expenditure 0.14 : 0.14/4 = 0.035 $\rightarrow$ 1 unit increase of the expenditure raises the winning probability by 3.5 percentage points **at most**

Odds and Odds Ratios

- Odds of success when the probability of success is p and that of failure is $1-p$:

$$\frac{p}{1-p}$$

  - $p = 0.5 \to$ odds is 1
  - $p = 1/3 \to$ odds is 0.5

- Odds ratio: the ratio of two sets of odds

$$\frac{p_1}{1-p_1} \Big/ \frac{p_2}{1-p_2}$$

- Merit of using odds ratios: no upper limit
- **Odds and odds ratios are different: do not confuse**

Interpreting Logistic Regression with Odds

- Odds of logistic regression:

$$\begin{aligned}\frac{\Pr(y=1|x)}{\Pr(y=0|x)} &= \left(\frac{\exp(\alpha+\beta x)}{1+\exp(\alpha+\beta x)}\right) \Big/ \left(1-\frac{\exp(\alpha+\beta x)}{1+\exp(\alpha+\beta x)}\right)\\ &= \exp(\alpha+\beta x)\end{aligned}$$

- Take natural logarithms of both sides

$$\log\left(\frac{\Pr(y=1|x)}{\Pr(y=0|x)}\right) = \alpha+\beta x = \operatorname{logit}(\alpha+\beta x)$$

- In logarithmic scale, 1 unit increase of x increase the odds by β
- In the original scale, the amount of change is $\exp(\beta)$
- E.g., $\exp(0.14) = 1.15$: 1 unit increase of $x \to$ the odds of winning is multiplied by 1.15: odds ratio is 1.15

Estimates and Standard Errors of Coefficients

- Purpose of logistic regression: estimate β in the linear predictor
- Estimation method: **maximum likelihood method**
- the point estimate of β
- $\hat{\beta} \pm 2\mathrm{se}$ are consistent with data
- Election example: $\hat{\beta} = 0.14$, se $= 0.01 \rightarrow \beta$ in $[0.14 \pm 2 \cdot 0.01] = [0.12, 0.16]$ are consistent with the data

Statistical Significance

- $\hat{\beta}$ is more than 2se away from 0: the effect is statistically significant
- Election example: $\hat{\beta} = 0.14$ is statistically significant positive effect: the electoral expenditure increases the probability of winning
- We don't discuss the significance of the intercept: we're not interested
- Careful interpretation is required for interaction terms

Let’s run some logistic regression with R

Next Week

Maximum Likelihood Method（最尤法）