Political Methodology III: Model Based Inference

Justin Grimmer

Associate Professor
Department of Political Science
Stanford University

April 13th, 2017


IN GOD WE TRUST

Going Public (Franco, Grimmer, and Whang 2017)

- Presidents interrupt prime time coverage $\rightsquigarrow$ why? what effect?
- There are at least 6 explanations (Canes Wrone 2001, 2006) (others: credit claim, veto bargain, beauty contest, highlight obstruction, stupid)
- Gathered serendipitous surveys$\rightsquigarrow$ happen to be in field when presidents go public (before$\rightsquigarrow$ control; after $\rightsquigarrow$ treatment) (we also have social media data and newspapers)
- Does going public increase probability respondents identify topic of president's speech as salient problem?
- Do respondents identify it as most important problem?

Modeling Bivariate Responses

$$
\begin{aligned}
Y_i &\in \{0, 1\} \\
X_{i1} &= \text{Treatment status (0/1)} \\
X_{i2} &= \text{Republican (0/1)}
\end{aligned}
$$

Infer effect of going public, condition on Republican as well

Modeling Bivariate Responses

$$
\begin{aligned}
Y_i &\in \{0,1\} \\
X_{i1} &= \text{ Treatment status (0/1)} \\
X_{i2} &= \text{Republican (0/1)}
\end{aligned}
$$

Infer effect of going public, condition on Republican as well

	Control	Treat
$\bar{Y}$	0.373	0.367

Linear Probabilty Model

$$
\begin{aligned}
Y_i &\sim \textsf{Normal}(\mu_i, \sigma^2) \\
E[Y_i|X_i, \boldsymbol{\beta}] &= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2}
\end{aligned}
$$

Linear Probabilty Model

$$\begin{aligned} Y_i &\sim \text{Normal}(\mu_i, \sigma^2) \\ E[Y_i | X_i, \boldsymbol{\beta}] &= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} \end{aligned}$$

$\beta_1 =$ Increase in the probability respondent i identifies problem as most important problem.

Linear Probabilty Model

$$
\begin{aligned}
Y_i &\sim \text{Normal}(\mu_i, \sigma^2) \\
E[Y_i | X_i, \boldsymbol{\beta}] &= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2}
\end{aligned}
$$

$\beta_1 =$ Increase in the probability respondent i identifies problem as most important problem.

Issues:

1) Predictions outside of 0 and 1
2) Constant effect
3) Efficiency loss

Probit Model

Suppose
$Y_i \sim$ Bernoulli(π_i)

Probit Model

Suppose
$Y_i \sim$ Bernoulli(π_i)
Assume:

$$\tilde{Y}_i \sim \text{Normal}(\mu_i, 1)$$

Probit Model

Suppose
$Y_i \sim$ Bernoulli(π_i)
Assume:

$$\begin{aligned}\tilde{Y}_i &\sim \text{Normal}(\mu_i, 1)\\ \tilde{Y}_i &= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i\end{aligned}$$

Probit Model

Suppose
$Y_i \sim \text{Bernoulli}(\pi_i)$
Assume:

$$
\begin{aligned}
\tilde{Y}_i &\sim \text{Normal}(\mu_i, 1) \\
\tilde{Y}_i &= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i \\
\epsilon_i &\sim \text{Normal}(0, 1)
\end{aligned}
$$

Probit Model

Suppose
$Y_i \sim \text{Bernoulli}(\pi_i)$
Assume:

$$\begin{aligned}
\tilde{Y}_i &\sim \text{Normal}(\mu_i, 1) \\
\tilde{Y}_i &= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i \\
\epsilon_i &\sim \text{Normal}(0, 1) \\
Y_i &= I(\tilde{Y}_i > 0)
\end{aligned}$$

Probit Model

Suppose
$Y_i \sim \text{Bernoulli}(\pi_i)$
Assume:

$$\begin{aligned}
\tilde{Y}_i &\sim \text{Normal}(\mu_i, 1) \\
\tilde{Y}_i &= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i \\
\epsilon_i &\sim \text{Normal}(0, 1) \\
Y_i &= I(\tilde{Y}_i > 0)
\end{aligned}$$

We will write:

Probit Model

Suppose
$Y_i \sim \text{Bernoulli}(\pi_i)$
Assume:

$$\begin{aligned}
\tilde{Y}_i &\sim \text{Normal}(\mu_i, 1) \\
\tilde{Y}_i &= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i \\
\epsilon_i &\sim \text{Normal}(0, 1) \\
Y_i &= I(\tilde{Y}_i > 0)
\end{aligned}$$

We will write:
$\boldsymbol{X}_i = (1, X_{i1}, X_{i2})$
$\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)$
$\boldsymbol{X}_i\boldsymbol{\beta} = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2}$

Probit Model

$$\pi_i = P(Y_i = 1) \quad = \quad P(\tilde{Y}_i > 0)$$

Probit Model

$$\begin{aligned}
\pi_i = P(Y_i = 1) &= P(\tilde{Y}_i > 0) \\
&= P(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i > 0)
\end{aligned}$$

Probit Model

$$
\begin{aligned}
\pi_i = P(Y_i = 1) &= P(\tilde{Y}_i > 0) \\
&= P(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i > 0) \\
&= P(\epsilon_i > -\boldsymbol{X}_i^{'}\boldsymbol{\beta})
\end{aligned}
$$

Probit Model

$$\begin{aligned}
\pi_i = P(Y_i = 1) &= P(\tilde{Y}_i > 0) \\
&= P(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i > 0) \\
&= P(\epsilon_i > -\boldsymbol{X}_i^{'}\boldsymbol{\beta}) \\
&= P(\epsilon_i < \boldsymbol{X}_i^{'}\boldsymbol{\beta})
\end{aligned}$$

Probit Model

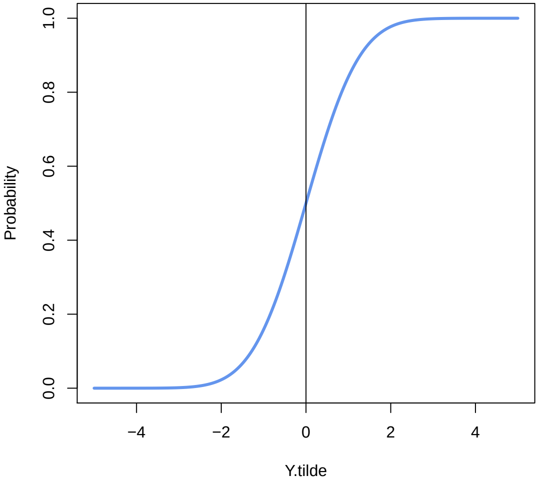
$$
\begin{aligned}
\pi_i = P(Y_i = 1) &= P(\tilde{Y}_i > 0) \\
&= P(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i > 0) \\
&= P(\epsilon_i > -\boldsymbol{X}_i^{'}\boldsymbol{\beta}) \\
&= P(\epsilon_i < \boldsymbol{X}_i^{'}\boldsymbol{\beta}) \\
&= \int_{-\infty}^{\boldsymbol{X}_i^{'}\boldsymbol{\beta}} \phi(\epsilon_i) d\epsilon_i
\end{aligned}
$$

Probit Model

$$
\begin{aligned}
\pi_i = P(Y_i = 1) &= P(\tilde{Y}_i > 0) \\
&= P(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i > 0) \\
&= P(\epsilon_i > -\boldsymbol{X}_i^{'}\boldsymbol{\beta}) \\
&= P(\epsilon_i < \boldsymbol{X}_i^{'}\boldsymbol{\beta}) \\
&= \int_{-\infty}^{\boldsymbol{X}_i^{'}\boldsymbol{\beta}} \phi(\epsilon_i) d\epsilon_i \\
&= \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})
\end{aligned}
$$

Probit Model

$$
\begin{aligned}
\pi_i = P(Y_i = 1) &= P(\tilde{Y}_i > 0) \\
&= P(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \epsilon_i > 0) \\
&= P(\epsilon_i > -\boldsymbol{X}_i^{'}\boldsymbol{\beta}) \\
&= P(\epsilon_i < \boldsymbol{X}_i^{'}\boldsymbol{\beta}) \\
&= \int_{-\infty}^{\boldsymbol{X}_i^{'}\boldsymbol{\beta}} \phi(\epsilon_i) d\epsilon_i \\
&= \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})
\end{aligned}
$$



1.0
0.8
0.6
0.4
0.2
0.0
Probability
−4
−2
0
2
4
Y.tilde

Probit Model

$$\begin{aligned} Y_i &\sim \text{Bernoulli}(\pi_i) \\ \pi_i &= \Phi(\boldsymbol{X}^{'}\boldsymbol{\beta}) \end{aligned}$$

Probit Model

$$\begin{aligned} Y_i &\sim \text{Bernoulli}(\pi_i) \\ \pi_i &= \Phi(\boldsymbol{X}^{'}\boldsymbol{\beta}) \end{aligned}$$

Then our likelihood:

Probit Model

$$\begin{aligned} Y_i &\sim \text{Bernoulli}(\pi_i) \\ \pi_i &= \Phi(\boldsymbol{X}^{'}\boldsymbol{\beta}) \end{aligned}$$

Then our likelihood:

$$L(\boldsymbol{\beta}|\boldsymbol{X}, \boldsymbol{Y}) = f(\boldsymbol{Y}|\boldsymbol{X}, \boldsymbol{\beta})$$

Probit Model

$$\begin{aligned} Y_i &\sim \text{Bernoulli}(\pi_i) \\ \pi_i &= \Phi(\boldsymbol{X}^{'}\boldsymbol{\beta}) \end{aligned}$$

Then our likelihood:

$$\begin{aligned} L(\boldsymbol{\beta}|\boldsymbol{X}, \boldsymbol{Y}) &= f(\boldsymbol{Y}|\boldsymbol{X}, \boldsymbol{\beta}) \\ &= \prod_{i=1}^{N} f(Y_i|\boldsymbol{X}_i\boldsymbol{\beta}) \end{aligned}$$

Probit Model

$$\begin{aligned} Y_i &\sim \text{Bernoulli}(\pi_i) \\ \pi_i &= \Phi(\boldsymbol{X}^{'}\boldsymbol{\beta}) \end{aligned}$$

Then our likelihood:

$$\begin{aligned} L(\boldsymbol{\beta}|\boldsymbol{X},\boldsymbol{Y}) &= f(\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{\beta}) \\ &= \prod_{i=1}^{N} f(Y_i|\boldsymbol{X}_i\boldsymbol{\beta}) \\ &= \prod_{i=1}^{N} \pi_i^{Y_i}(1-\pi_i)^{1-Y_i} \end{aligned}$$

Probit Model

$$\begin{aligned} Y_i &\sim \text{Bernoulli}(\pi_i) \\ \pi_i &= \Phi(\boldsymbol{X}^{'}\boldsymbol{\beta}) \end{aligned}$$

Then our likelihood:

$$\begin{aligned} L(\boldsymbol{\beta}|\boldsymbol{X},\boldsymbol{Y}) &= f(\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{\beta}) \\ &= \prod_{i=1}^{N} f(Y_i|\boldsymbol{X}_i\boldsymbol{\beta}) \\ &= \prod_{i=1}^{N} \pi_i^{Y_i}(1-\pi_i)^{1-Y_i} \\ &= \prod_{i=1}^{N} \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})^{Y_i}(1-\Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}))^{1-Y_i} \end{aligned}$$

Probit Model

$$L(\boldsymbol{\beta}|\boldsymbol{X}, \boldsymbol{Y}) = \prod_{i=1}^{N} \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})^{Y_i}(1 - \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}))^{1-Y_i}$$

Probit Model

$$
\begin{aligned}
L(\boldsymbol{\beta}|\boldsymbol{X}, \boldsymbol{Y}) &= \prod_{i=1}^{N} \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})^{Y_i}(1 - \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}))^{1-Y_i} \\
\log L(\boldsymbol{\beta}|\boldsymbol{X}, \boldsymbol{Y}) &= \sum_{i=1}^{N} \Big(Y_i \log \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}) + (1 - Y_i) \log(1 - \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})) \Big)
\end{aligned}
$$

Probit Model

$$
\begin{aligned}
L(\boldsymbol{\beta}|\boldsymbol{X}, \boldsymbol{Y}) & = \prod_{i=1}^{N} \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})^{Y_i}(1 - \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}))^{1-Y_i} \\
\log L(\boldsymbol{\beta}|\boldsymbol{X}, \boldsymbol{Y}) & = \sum_{i=1}^{N} \left(Y_i \log \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}) + (1 - Y_i) \log(1 - \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}))\right)
\end{aligned}
$$

Maximize likelihood with respect to $\boldsymbol{\beta}$

Probit Model

$$l(\boldsymbol{\beta}|\boldsymbol{X}, \boldsymbol{Y}) \quad = \quad \sum_{i=1}^{N} \Big(Y_i \log \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}) + (1 - Y_i) \log(1 - \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})) \Big)$$

Probit Model

$$l(\boldsymbol{\beta}|\boldsymbol{X},\boldsymbol{Y}) \;=\; \sum_{i=1}^{N}\Big(Y_i \log \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}) + (1-Y_i)\log(1-\Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}))\Big)$$

Score Function for β_k:

$$\frac{\partial l(\boldsymbol{\beta}|\boldsymbol{X}_i, Y_i)}{\partial \beta_k} \;=\; \left(Y_i \frac{\phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})}{\Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})} - (1-Y_i)\frac{\phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})}{1-\Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})}\right) X_k$$

Probit Model

$$l(\boldsymbol{\beta}|\boldsymbol{X},\boldsymbol{Y}) \;=\; \sum_{i=1}^{N}\Big(Y_i \log \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}) + (1-Y_i)\log(1-\Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}))\Big)$$

Score Function for β_k:

$$\frac{\partial l(\boldsymbol{\beta}|\boldsymbol{X}_i, Y_i)}{\partial \beta_k} \;=\; \left(Y_i \frac{\phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})}{\Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})} - (1-Y_i)\frac{\phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})}{1-\Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})}\right) X_k$$

Hessian (typical entry h_{kj}):

$$\begin{aligned}\frac{\partial^2 l(\boldsymbol{\beta}|\boldsymbol{X}_i, Y_i)}{\partial \beta_k \partial \beta_j} \;&=\; \phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})[Y_i \frac{\phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}) + \boldsymbol{X}_i\boldsymbol{\beta}\Phi(\boldsymbol{X}_i\boldsymbol{\beta})}{\Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta})^2} \\ &\quad + (1-Y_i)\frac{\phi(\boldsymbol{X}_i\boldsymbol{\beta}) - \boldsymbol{X}_i^{'}\boldsymbol{\beta}(1-\Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}))}{(1-\Phi(\boldsymbol{X}_i\boldsymbol{\beta}))^2}] X_{ik}X_{ij}\end{aligned}$$

Probit Model

Maximum Likelihood Estimates

$$0 = \sum_{i=1}^{N} \left(Y_i \frac{\phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}^*)}{\Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}^*)} - (1 - Y_i)\frac{\phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}^*)}{1 - \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}^*)} \right) X_k$$

Solve for $\boldsymbol{\beta}$?

Probit Model

Maximum Likelihood Estimates

$$0 = \sum_{i=1}^{N} \left(Y_i \frac{\phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}^*)}{\Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}^*)} - (1 - Y_i) \frac{\phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}^*)}{1 - \Phi(\boldsymbol{X}_i^{'}\boldsymbol{\beta}^*)} \right) X_k$$

Solve for $\boldsymbol{\beta}$? Computational Approaches

- Newton Raphson
- BFGS (Quasi-Newton. Requires only an approximate Hessian.)

Probit Model: Maximum Likelihood Estimates

R Code

Probit Model: Most Important Problem

	Most Important Problem
Intercept	-0.26
	(0.03)
Post-Speech	-0.01
	(0.05)
Republican	-0.20
	(0.05)

How do we interpret the coefficients?

Probit Model: Most Important Problem

	Most Important Problem
Intercept	-0.26
	(0.03)
Post-Speech	-0.01
	(0.05)
Republican	-0.20
	(0.05)

How do we interpret the coefficients?

They are on the latent scale $\rightsquigarrow$ Need to define all values when determining predicted probabilities

Probit Model: Interpreting the Coefficients

Expected Value:

Probit Model: Interpreting the Coefficients

Expected Value:

$$E[Y_i|\boldsymbol{X}_i] = \Phi(\beta_0^* + \beta_1^* X_{i1} + \beta_2^* X_{i2})$$

Probit Model: Interpreting the Coefficients

Expected Value:

$$
\begin{aligned}
E[Y_i|\boldsymbol{X}_i] &= \Phi(\beta_0^* + \beta_1^* X_{i1} + \beta_2^* X_{i2}) \\
E[Y_i|X_{i1} = 0, X_{i2} = 0] &= \Phi(-0.26 - 0.01 \times 0 - 0.20 \times 0) = 0.40
\end{aligned}
$$

Probit Model: Interpreting the Coefficients

Expected Value:

$$
\begin{aligned}
E[Y_i|\boldsymbol{X}_i] &= \Phi(\beta_0^* + \beta_1^* X_{i1} + \beta_2^* X_{i2}) \\
E[Y_i|X_{i1} = 0, X_{i2} = 0] &= \Phi(-0.26 - 0.01 \times 0 - 0.20 \times 0) = 0.40 \\
E[Y_i|X_{i1} = 0, X_{i2} = 1] &= \Phi(-0.26 - 0.01 \times 0 - 0.20 \times 1) = 0.32
\end{aligned}
$$

Probit Model: Interpreting the Coefficients

Expected Value:

$$\begin{aligned} E[Y_i|\boldsymbol{X}_i] &= \Phi(\beta_0^* + \beta_1^* X_{i1} + \beta_2^* X_{i2}) \\ E[Y_i|X_{i1} = 0, X_{i2} = 0] &= \Phi(-0.26 - 0.01 \times 0 - 0.20 \times 0) = 0.40 \\ E[Y_i|X_{i1} = 0, X_{i2} = 1] &= \Phi(-0.26 - 0.01 \times 0 - 0.20 \times 1) = 0.32 \end{aligned}$$

First difference:

Probit Model: Interpreting the Coefficients

Expected Value:

$$
\begin{aligned}
E[Y_i|\boldsymbol{X}_i] &= \Phi(\beta_0^* + \beta_1^* X_{i1} + \beta_2^* X_{i2}) \\
E[Y_i|X_{i1} = 0, X_{i2} = 0] &= \Phi(-0.26 - 0.01 \times 0 - 0.20 \times 0) = 0.40 \\
E[Y_i|X_{i1} = 0, X_{i2} = 1] &= \Phi(-0.26 - 0.01 \times 0 - 0.20 \times 1) = 0.32
\end{aligned}
$$

First difference:

$$
\begin{aligned}
E[Y_i|X_{i1} = 1, X_{i2} = 0] - E[Y_i|X_{i1} = 0, X_{i2} = 0] &= \\
\Phi(-0.26 - 0.01) - \Phi(-0.26) &= -0.0039
\end{aligned}
$$

Probit Model: Interpreting the Coefficients

Expected Value:

$$
\begin{aligned}
E[Y_i|\boldsymbol{X}_i] &= \Phi(\beta_0^* + \beta_1^* X_{i1} + \beta_2^* X_{i2}) \\
E[Y_i|X_{i1} = 0, X_{i2} = 0] &= \Phi(-0.26 - 0.01 \times 0 - 0.20 \times 0) = 0.40 \\
E[Y_i|X_{i1} = 0, X_{i2} = 1] &= \Phi(-0.26 - 0.01 \times 0 - 0.20 \times 1) = 0.32
\end{aligned}
$$

First difference:

$$
\begin{aligned}
E[Y_i|X_{i1} = 1, X_{i2} = 0] - E[Y_i|X_{i1} = 0, X_{i2} = 0] &= \\
\Phi(-0.26 - 0.01) - \Phi(-0.26) &= -0.0039 \\
E[Y_i|X_{i1} = 1, X_{i2} = 1] - E[Y_i|X_{i1} = 0, X_{i2} = 1] &= \\
\Phi(-0.26 - 0.01 - 0.2) - \Phi(-0.26 - 0.2) &= -0.0036
\end{aligned}
$$

Probit Model: Interpreting the Coefficients

Expected Value:

$$\begin{aligned}
E[Y_i|\boldsymbol{X}_i] &= \Phi(\beta_0^* + \beta_1^* X_{i1} + \beta_2^* X_{i2}) \\
E[Y_i|X_{i1}=0, X_{i2}=0] &= \Phi(-0.26 - 0.01 \times 0 - 0.20 \times 0) = 0.40 \\
E[Y_i|X_{i1}=0, X_{i2}=1] &= \Phi(-0.26 - 0.01 \times 0 - 0.20 \times 1) = 0.32
\end{aligned}$$

First difference:

$$\begin{aligned}
E[Y_i|X_{i1}=1, X_{i2}=0] - E[Y_i|X_{i1}=0, X_{i2}=0] &= \\
\Phi(-0.26 - 0.01) - \Phi(-0.26) &= -0.0039 \\
E[Y_i|X_{i1}=1, X_{i2}=1] - E[Y_i|X_{i1}=0, X_{i2}=1] &= \\
\Phi(-0.26 - 0.01 - 0.2) - \Phi(-0.26 - 0.2) &= -0.0036
\end{aligned}$$

We will perform inference on these on Monday!

Probit Model: Interpreting the Coefficients

$$\begin{aligned}
\frac{\partial \Phi(\boldsymbol{X}_i \boldsymbol{\beta})}{\partial X_{ij}} &= \phi(\boldsymbol{X}_i \beta)\beta_j \\
\textsf{Max Effect} &= 0.4 \times \beta_j
\end{aligned}$$

Logit Model

Define:

$$
\begin{aligned}
\mathsf{Odds}(\pi) &= \frac{\pi}{1-\pi} \\
\log \mathsf{Odds}(\pi) &= \log\left(\frac{\pi}{1-\pi}\right) \\
\mathsf{logit}(\pi) &\equiv \log\left(\frac{\pi}{1-\pi}\right)
\end{aligned}
$$

Logit Model

$$\log\left(\frac{\pi}{1-\pi}\right) = \alpha$$

Logit Model

$$
\begin{aligned}
\log\left(\frac{\pi}{1-\pi}\right) &= \alpha \\
\left(\frac{\pi}{1-\pi}\right) &= \exp(\alpha)
\end{aligned}
$$

Logit Model

$$
\begin{aligned}
\log\left(\frac{\pi}{1-\pi}\right) &= \alpha \\
\left(\frac{\pi}{1-\pi}\right) &= \exp(\alpha) \\
\pi &= \exp(\alpha)(1-\pi)
\end{aligned}
$$

Logit Model

$$
\begin{aligned}
\log\left(\frac{\pi}{1-\pi}\right) &= \alpha \\
\left(\frac{\pi}{1-\pi}\right) &= \exp(\alpha) \\
\pi &= \exp(\alpha)(1-\pi) \\
\pi(1+\exp(\alpha)) &= \exp(\alpha)
\end{aligned}
$$

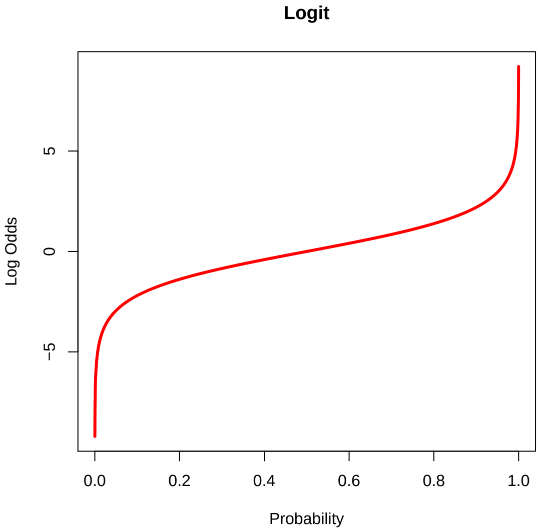
Logit Model

$$
\begin{aligned}
\log\left(\frac{\pi}{1-\pi}\right) &= \alpha \\
\left(\frac{\pi}{1-\pi}\right) &= \exp(\alpha) \\
\pi &= \exp(\alpha)(1-\pi) \\
\pi(1+\exp(\alpha)) &= \exp(\alpha) \\
\pi &= \frac{\exp(\alpha)}{1+\exp(\alpha)} = \mathsf{Logit}^{-1}(\alpha)
\end{aligned}
$$

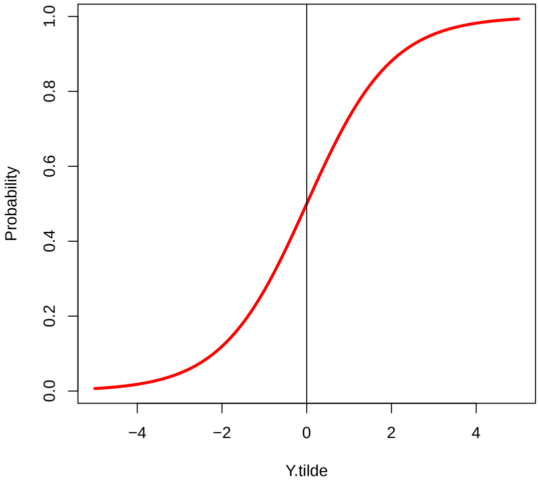
Logit Model

$$
\begin{aligned}
\log\left(\frac{\pi}{1-\pi}\right) &= \alpha \\
\left(\frac{\pi}{1-\pi}\right) &= \exp(\alpha) \\
\pi &= \exp(\alpha)(1-\pi) \\
\pi(1+\exp(\alpha)) &= \exp(\alpha) \\
\pi &= \frac{\exp(\alpha)}{1+\exp(\alpha)} = \textsf{Logit}^{-1}(\alpha) \\
\frac{\exp(\alpha)}{1+\exp(\alpha)} &= \frac{1}{1+\exp(-\alpha)} = \textsf{Logistic}(\alpha)
\end{aligned}
$$

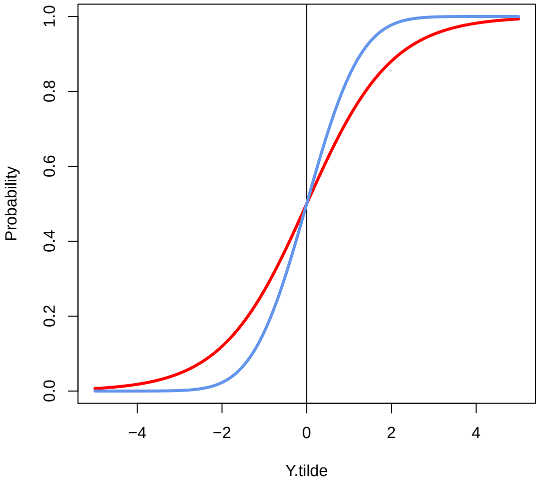
Logit Model



Logit Model



Logit Model



$$Z \sim \text{Logistic}(\mu, s)$$

$$
\begin{aligned}
Z &\sim \mathsf{Logistic}(\mu, s) \\
f(z) &= \frac{\exp(\frac{z-\mu}{s})}{s(1+\exp(\frac{z-\mu}{s}))^2}
\end{aligned}
$$

$$
\begin{aligned}
Z &\sim \mathsf{Logistic}(\mu, s) \\
f(z) &= \frac{\exp(\frac{z-\mu}{s})}{s(1+\exp(\frac{z-\mu}{s}))^2} \\
Z &\sim \underbrace{\mathsf{Logistic}(0,1)}_{\text{Standard Logistic}}
\end{aligned}
$$

$$
\begin{aligned}
Z &\sim \text{Logistic}(\mu, s) \\
f(z) &= \frac{\exp(\frac{z-\mu}{s})}{s(1+\exp(\frac{z-\mu}{s}))^2} \\
Z &\sim \underbrace{\text{Logistic}(0,1)}_{\text{Standard Logistic}} \\
f(z) &= \frac{\exp(z)}{(1+\exp(z))^2}
\end{aligned}
$$

Logit Model

$$\begin{aligned}\tilde{Y}_i &= \boldsymbol{X}_i^{'}\boldsymbol{\beta} + \epsilon_i \\ \epsilon_i &\sim \textsf{Logistic}(0, 1)\end{aligned}$$

Logit Model

$$\begin{aligned}
\tilde{Y}_i &= \boldsymbol{X}_i^{'}\boldsymbol{\beta} + \epsilon_i \\
\epsilon_i &\sim \textsf{Logistic}(0, 1) \\
P(Y_i = 1|\boldsymbol{X}_i, \boldsymbol{\beta}) &= I(\tilde{Y}_i > 0)
\end{aligned}$$

Logit Model

$$\begin{aligned}
\tilde{Y}_i &= \boldsymbol{X}_i^{'}\boldsymbol{\beta} + \epsilon_i \\
\epsilon_i &\sim \textsf{Logistic}(0,1) \\
P(Y_i = 1|\boldsymbol{X}_i, \boldsymbol{\beta}) &= I(\tilde{Y}_i > 0) \\
&= I(\boldsymbol{X}_i^{'}\boldsymbol{\beta} + \epsilon_i > 0)
\end{aligned}$$

Logit Model

$$
\begin{aligned}
\tilde{Y}_i &= \boldsymbol{X}_i^{'}\boldsymbol{\beta} + \epsilon_i \\
\epsilon_i &\sim \textsf{Logistic}(0, 1) \\
P(Y_i = 1|\boldsymbol{X}_i, \boldsymbol{\beta}) &= I(\tilde{Y}_i > 0) \\
&= I(\boldsymbol{X}_i^{'}\boldsymbol{\beta} + \epsilon_i > 0) \\
&= I(\epsilon_i > -\boldsymbol{X}_i^{'}\boldsymbol{\beta})
\end{aligned}
$$

Logit Model

$$\begin{aligned}
\tilde{Y}_i &= \boldsymbol{X}_i^{'}\boldsymbol{\beta} + \epsilon_i \\
\epsilon_i &\sim \textsf{Logistic}(0,1) \\
P(Y_i = 1 | \boldsymbol{X}_i, \boldsymbol{\beta}) &= I(\tilde{Y}_i > 0) \\
&= I(\boldsymbol{X}_i^{'}\boldsymbol{\beta} + \epsilon_i > 0) \\
&= I(\epsilon_i > -\boldsymbol{X}_i^{'}\boldsymbol{\beta}) \\
&= I(\epsilon_i < \boldsymbol{X}_i^{'}\boldsymbol{\beta})
\end{aligned}$$

Logit Model

$$
\begin{aligned}
\tilde{Y}_i &= \boldsymbol{X}_i^{'}\boldsymbol{\beta} + \epsilon_i \\
\epsilon_i &\sim \textsf{Logistic}(0,1) \\
P(Y_i = 1|\boldsymbol{X}_i, \boldsymbol{\beta}) &= I(\tilde{Y}_i > 0) \\
&= I(\boldsymbol{X}_i^{'}\boldsymbol{\beta} + \epsilon_i > 0) \\
&= I(\epsilon_i > -\boldsymbol{X}_i^{'}\boldsymbol{\beta}) \\
&= I(\epsilon_i < \boldsymbol{X}_i^{'}\boldsymbol{\beta}) \\
&= \textsf{Logistic}(\boldsymbol{X}_i^{'}\boldsymbol{\beta}) = F(\boldsymbol{X}_i^{'}\boldsymbol{\beta})
\end{aligned}
$$

Logit Model

$$
\begin{aligned}
Y_i &\sim \text{Bernoulli}(\pi_i) \\
\pi_i &= F(\boldsymbol{X}_i^{'}\boldsymbol{\beta})
\end{aligned}
$$

Likelihood function

$$
\begin{aligned}
L(\boldsymbol{\beta}|\boldsymbol{X}, \boldsymbol{Y}) &= f(\boldsymbol{Y}|\boldsymbol{X}, \boldsymbol{\beta}) \\
&= \prod_{i=1}^{N} f(Y_i|\boldsymbol{X}_i, \boldsymbol{\beta}) \\
&= \prod_{i=1}^{N} F(\boldsymbol{X}_i^{'}\boldsymbol{\beta})^{Y_i}(1 - F(\boldsymbol{X}_i^{'}\boldsymbol{\beta}))^{1-Y_i}
\end{aligned}
$$

Logit Model

Log-likelihood:

$$l(\boldsymbol{\beta}|\boldsymbol{X}, \boldsymbol{Y}) \quad = \quad \sum_{i=1}^{N} Y_i \log F(\boldsymbol{X}_i^{'}\boldsymbol{\beta}) + (1 - Y_i) \log(1 - F(\boldsymbol{X}_i^{'}\boldsymbol{\beta}))$$

Homework: Derive Score + Hessian for Logistic Normal + Implement with Optim

R Code

Logit Model: Most Important Problem

Most Important Problem

	Probit	Logit
Intercept	-0.26	-0.41
	(0.03)	(0.05)
Post-Speech	-0.01	-0.02
	(0.05)	(0.09)
Republican	-0.20	-0.32
	(0.05)	(0.08)

Logit Model: Interpreting the Coefficients

Expected Value:

Logit Model: Interpreting the Coefficients

Expected Value:

$$E[Y_i|\boldsymbol{X}_i] = \frac{1}{1 + \exp(-\beta_0 - \beta_1 X_{i1} - \beta_2 X_{i2})}$$

Logit Model: Interpreting the Coefficients

Expected Value:

$$
\begin{aligned}
E[Y_i|\boldsymbol{X}_i] &= \frac{1}{1+\exp(-\beta_0-\beta_1 X_{i1}-\beta_2 X_{i2})} \\
E[Y_i|X_{i1}=0, X_{i2}=0] &= \frac{1}{1+\exp(0.41)} = 0.40
\end{aligned}
$$

Logit Model: Interpreting the Coefficients

Expected Value:

$$
\begin{aligned}
E[Y_i|\boldsymbol{X}_i] &= \frac{1}{1+\exp(-\beta_0-\beta_1 X_{i1}-\beta_2 X_{i2})} \\
E[Y_i|X_{i1}=0, X_{i2}=0] &= \frac{1}{1+\exp(0.41)} = 0.40 \\
E[Y_i|X_{i1}=0, X_{i2}=1] &= \frac{1}{1+\exp(0.41+0.32)} = 0.33
\end{aligned}
$$

Logit Model: Interpreting the Coefficients

Expected Value:

$$
\begin{aligned}
E[Y_i|\boldsymbol{X}_i] &= \frac{1}{1+\exp(-\beta_0-\beta_1 X_{i1}-\beta_2 X_{i2})} \\
E[Y_i|X_{i1}=0, X_{i2}=0] &= \frac{1}{1+\exp(0.41)} = 0.40 \\
E[Y_i|X_{i1}=0, X_{i2}=1] &= \frac{1}{1+\exp(0.41+0.32)} = 0.33
\end{aligned}
$$

First Differences

Logit Model: Interpreting the Coefficients

Expected Value:

$$
\begin{aligned}
E[Y_i|\boldsymbol{X}_i] &= \frac{1}{1+\exp(-\beta_0-\beta_1 X_{i1}-\beta_2 X_{i2})} \\
E[Y_i|X_{i1}=0, X_{i2}=0] &= \frac{1}{1+\exp(0.41)} = 0.40 \\
E[Y_i|X_{i1}=0, X_{i2}=1] &= \frac{1}{1+\exp(0.41+0.32)} = 0.33
\end{aligned}
$$

First Differences

$$
\begin{aligned}
E[Y_i|X_{i1}=1, X_{i2}=0] - E[Y_i|X_{i1}=0, X_{i2}=0] &= \\
0.3941 - 0.3989 &= -0.005
\end{aligned}
$$

Logit Model: Interpreting the Coefficients

Expected Value:

$$\begin{aligned}
E[Y_i|\boldsymbol{X}_i] &= \frac{1}{1+\exp(-\beta_0-\beta_1 X_{i1}-\beta_2 X_{i2})} \\
E[Y_i|X_{i1}=0, X_{i2}=0] &= \frac{1}{1+\exp(0.41)} = 0.40 \\
E[Y_i|X_{i1}=0, X_{i2}=1] &= \frac{1}{1+\exp(0.41+0.32)} = 0.33
\end{aligned}$$

First Differences

$$\begin{aligned}
E[Y_i|X_{i1}=1, X_{i2}=0] - E[Y_i|X_{i1}=0, X_{i2}=0] &= \\
0.3941 - 0.3989 &= -0.005 \\
E[Y_i|X_{i1}=1, X_{i2}=1] - E[Y_i|X_{i1}=0, X_{i2}=1] &= \\
0.3208 - 0.32519 &= -0.0043
\end{aligned}$$

Logit Model: Interpreting the Coefficients

$$
\begin{aligned}
\frac{\partial F(\boldsymbol{X}_i^{'}\boldsymbol{\beta})}{\partial X_{ij}} &= f(\boldsymbol{X}_i^{'}\boldsymbol{\beta})\beta_j \\
\text{Max Effect} &= 0.25 \times \beta_j
\end{aligned}
$$

Most Important Problem: Linear Probability Model

	Probit	Logit	LPM
Intercept	-0.26	-0.41	0.40
	(0.03)	(0.05)	(0.13)
Post-Speech	-0.01	-0.02	-0.004
	(0.05)	(0.09)	(0.02)
Republican	-0.20	-0.32	-0.07
	(0.05)	(0.08)	(0.02)

LPM might (usually) be ok, but

1) Probit/Logit = easy to run
2) Base for more complicated models