

# PLSC 503 – Spring 2020

## Binary Response Models, I

April 7, 2020

# Linear Probability Model (LPM)

$$E(Y) = \mathbf{X}\beta$$

$$Y \in \{0, 1\}$$

$$\begin{aligned} E(Y) &= 1[\Pr(Y = 1)] + 0[\Pr(Y = 0)] \\ &= \Pr(Y = 1) \end{aligned}$$

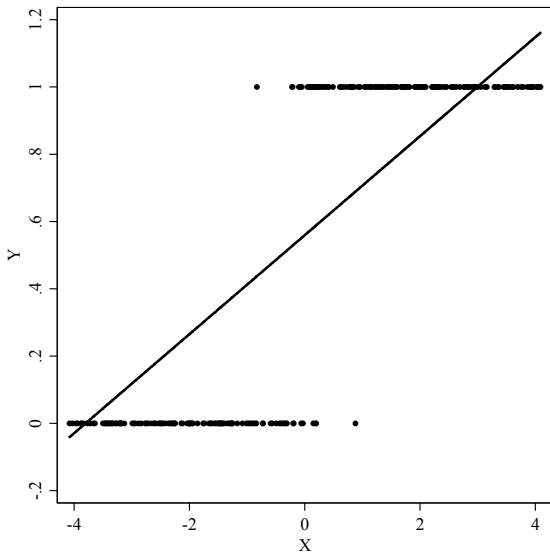
So:

$$\Pr(Y_i = 1) = \mathbf{X}_i\beta$$

or:

$$Y_i = \mathbf{X}_i\beta + u_i$$

# LPM Illustrated



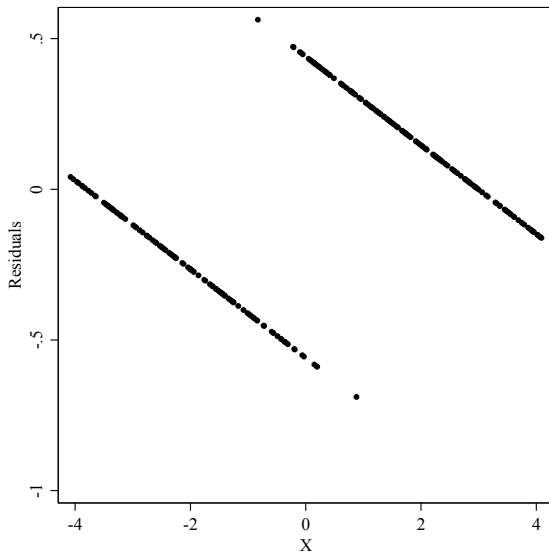
Variance:

$$\begin{aligned}\text{Var}(Y) &= E(Y)[1 - E(Y)] \\ &= \mathbf{X}_i\boldsymbol{\beta}(1 - \mathbf{X}_i\boldsymbol{\beta})\end{aligned}$$

Residuals:

$$\hat{u}_i \in \{1 - \mathbf{X}_i\hat{\boldsymbol{\beta}}, -\mathbf{X}_i\hat{\boldsymbol{\beta}}\}$$

# LPM Residuals



## LPM Issues (continued)

- Predictions  $\notin [0, 1]$
- Functional form  $\rightarrow \frac{\partial E(Y)}{\partial X} = \beta$  (reasonable?)

## A Different Model

$$Y_i^* = \mathbf{X}_i\boldsymbol{\beta} + u_i$$

$$Y_i = 0 \text{ if } Y_i^* < 0$$

$$Y_i = 1 \text{ if } Y_i^* \geq 0$$

So:

$$\begin{aligned}\Pr(Y_i = 1) &= \Pr(Y_i^* \geq 0) \\ &= \Pr(\mathbf{X}_i\boldsymbol{\beta} + u_i \geq 0) \\ &= \Pr(u_i \geq -\mathbf{X}_i\boldsymbol{\beta}) \\ &= \Pr(u_i \leq \mathbf{X}_i\boldsymbol{\beta}) \\ &= \int_{-\infty}^{\mathbf{X}_i\boldsymbol{\beta}} f(u) du\end{aligned}$$

“Standard logistic” PDF:

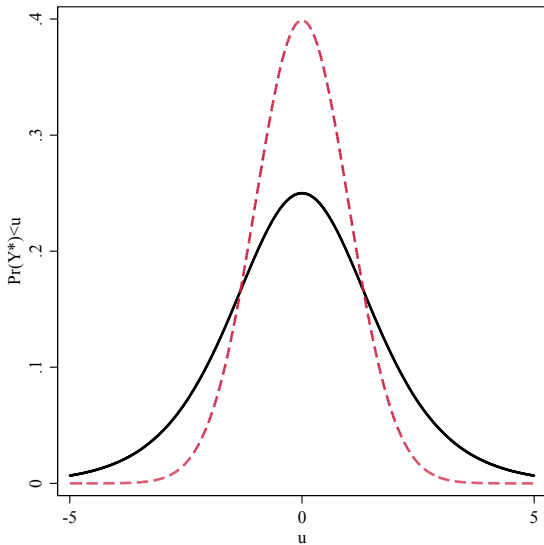
$$\Pr(u) \equiv \lambda(u) = \frac{\exp(u)}{[1 + \exp(u)]^2}$$

CDF:

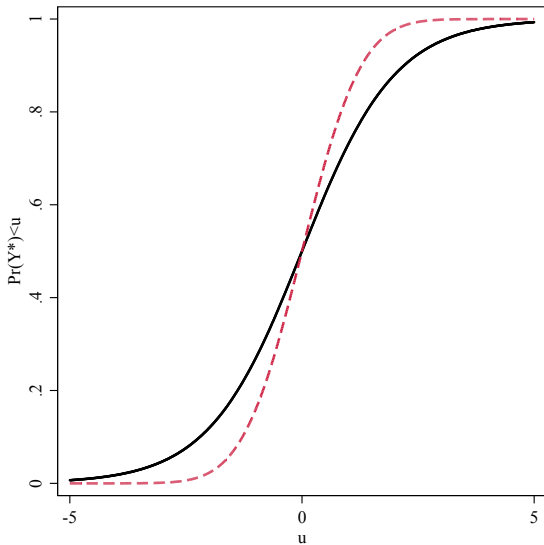
$$\begin{aligned}\Lambda(u) &= \int \lambda(u) du \\ &= \frac{\exp(u)}{1 + \exp(u)} \\ &= \frac{1}{1 + \exp(-u)}\end{aligned}$$



# Standard Normal and Logistic PDFs



# Standard Normal and Logistic CDFs



- $\lambda(u) = 1 - \lambda(-u)$
- $\Lambda(u) = 1 - \Lambda(-u)$
- $\text{Var}(u) = \frac{\pi^2}{3} \approx 3.29$

## Logistic $\rightarrow$ “Logit”

$$\begin{aligned}\Pr(Y_i = 1) &= \Pr(Y_i^* > 0) \\ &= \Pr(u_i \leq \mathbf{X}_i\boldsymbol{\beta}) \\ &= \Lambda(\mathbf{X}_i\boldsymbol{\beta}) \\ &= \frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1 + \exp(\mathbf{X}_i\boldsymbol{\beta})}\end{aligned}$$

$$\text{(equivalently)} = \frac{1}{1 + \exp(-\mathbf{X}_i\boldsymbol{\beta})}$$

$$L_i = \left( \frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1 + \exp(\mathbf{X}_i\boldsymbol{\beta})} \right)^{Y_i} \left[ 1 - \left( \frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1 + \exp(\mathbf{X}_i\boldsymbol{\beta})} \right) \right]^{1-Y_i}$$

$$L = \prod_{i=1}^N \left( \frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1 + \exp(\mathbf{X}_i\boldsymbol{\beta})} \right)^{Y_i} \left[ 1 - \left( \frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1 + \exp(\mathbf{X}_i\boldsymbol{\beta})} \right) \right]^{1-Y_i}$$

$$\begin{aligned} \ln L = & \sum_{i=1}^N Y_i \ln \left( \frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1 + \exp(\mathbf{X}_i\boldsymbol{\beta})} \right) + \\ & (1 - Y_i) \ln \left[ 1 - \left( \frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1 + \exp(\mathbf{X}_i\boldsymbol{\beta})} \right) \right] \end{aligned}$$

# Be Normal?

$$\Pr(u) \equiv \phi(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)$$

$$\Phi(u) = \int_{-\infty}^u \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du$$

Normal  $\rightarrow$  “Probit”

$$\begin{aligned}\Pr(Y_i = 1) &= \Phi(\mathbf{X}_i\boldsymbol{\beta}) \\ &= \int_{-\infty}^{\mathbf{X}_i\boldsymbol{\beta}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(\mathbf{X}_i\boldsymbol{\beta})^2}{2}\right) d\mathbf{X}_i\boldsymbol{\beta}\end{aligned}$$

$$L = \prod_{i=1}^N [\Phi(\mathbf{X}_i\boldsymbol{\beta})]^{Y_i} [1 - \Phi(\mathbf{X}_i\boldsymbol{\beta})]^{(1-Y_i)}$$

$$\ln L = \sum_{i=1}^N Y_i \ln \Phi(\mathbf{X}_i\boldsymbol{\beta}) + (1 - Y_i) \ln [1 - \Phi(\mathbf{X}_i\boldsymbol{\beta})]$$

# Digression I: Logit as an Odds Model

$$\text{Odds}(Z) \equiv \Omega(Z) = \frac{\Pr(Z)}{1 - \Pr(Z)}.$$

$$\ln[\Omega(Z)] = \ln \left[ \frac{\Pr(Z)}{1 - \Pr(Z)} \right]$$

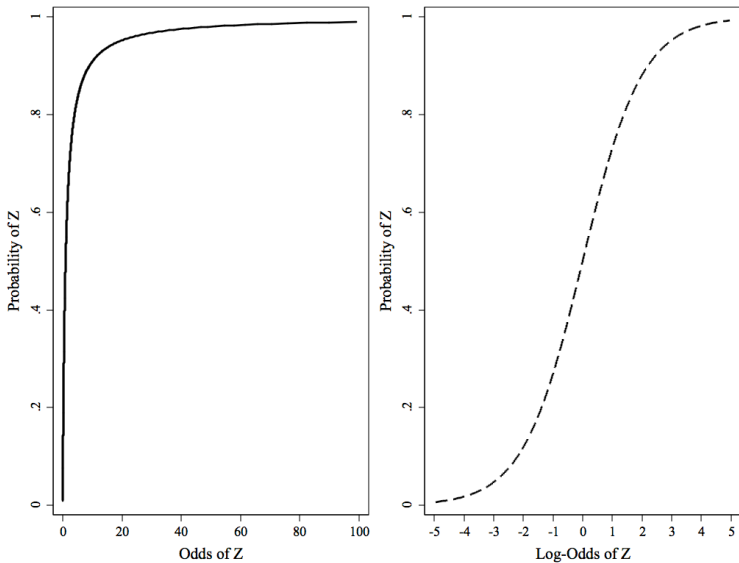
$$\ln[\Omega(Z_i)] = \mathbf{X}_i\beta$$

$$\begin{aligned}\Omega(Z_i) &= \frac{\Pr(Z)}{1 - \Pr(Z)} \\ &= \exp(\mathbf{X}_i\beta)\end{aligned}$$

$$\Pr(Z_i) = \frac{\exp(\mathbf{X}_i\beta)}{1 + \exp(\mathbf{X}_i\beta)}$$



# Visualizing Log-Odds



## Digression II: The Random Utility Model

$$Y \in \{SQ, A\}$$

$$\begin{aligned} Y_i &= A \quad \text{if} \quad E[U_i(A)] \geq E[U_i(SQ)] \\ &= SQ \quad \text{if} \quad E[U_i(A)] < E[U_i(SQ)] \end{aligned}$$

$$E[U_i(A)] = \mathbf{X}_{iA}\boldsymbol{\beta} + u_{iA}$$

So:

$$\begin{aligned} \Pr(Y = A) &= \Pr\{E[U_i(A)] \geq E[U_i(SQ)]\} \\ &= \Pr\{(\mathbf{X}_{iA}\boldsymbol{\beta} + u_{iA}) \geq E[U_i(SQ)]\} \end{aligned}$$

## Digression II: The Random Utility Model

Normalize:

$$E[U_i(SQ)] = 0$$

Then:

$$\begin{aligned}\Pr(Y = A) &= \Pr\{(\mathbf{X}_{iA}\boldsymbol{\beta} + u_{iA}) \geq 0\} \\ &= \Pr\{u_{iA} \geq -\mathbf{X}_{iA}\boldsymbol{\beta}\} \\ &= F(\mathbf{X}_{iA}\boldsymbol{\beta})\end{aligned}$$

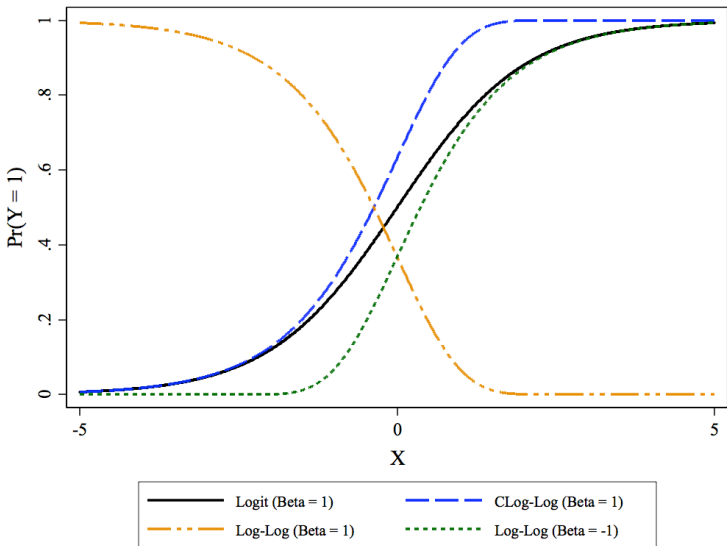
## Other Models: Complementary Log-Log

$$\Pr(Y_i = 1) = 1 - \exp[-\exp(\mathbf{X}_i\beta)]$$

or

$$\ln\{-\ln[1 - \Pr(Y_i = 1)]\} = \mathbf{X}_i\beta$$

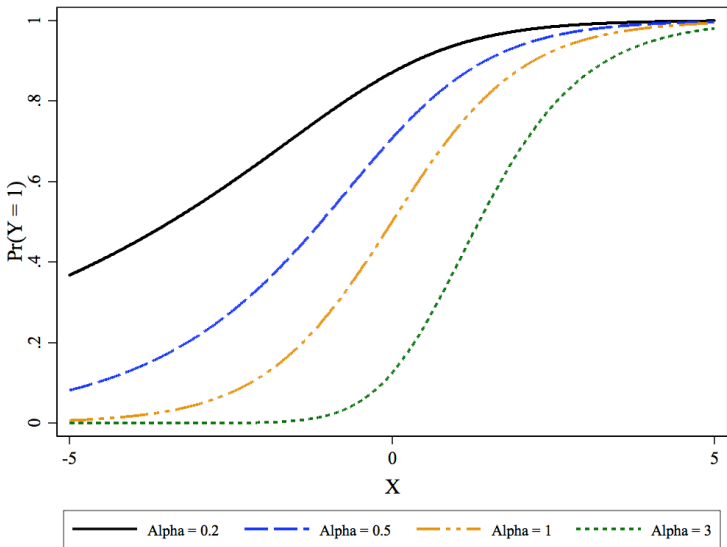
# Logit and C-log-log CDFs



$$\Pr(Y_i = 1) = \frac{1}{[1 + \exp(-\mathbf{X}_i\boldsymbol{\beta})]^\alpha}, \quad \alpha > 0$$

$$\begin{aligned} \alpha = 1 \rightarrow \frac{1}{[1 + \exp(-\mathbf{X}_i\boldsymbol{\beta})]^1} &= \frac{1}{1 + \exp(-\mathbf{X}_i\boldsymbol{\beta})} \\ &= \frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1 + \exp(\mathbf{X}_i\boldsymbol{\beta})} \end{aligned}$$

# Scobit, Visualized



# Binary Response Models: Identification

- “Threshold” =  $Y^* > 0$
- $E(u_i | \mathbf{X}, \beta) = 0$
- $\text{Var}(u_i) = \frac{\pi^2}{3}$  or 1.0.



# Logit vs. Probit

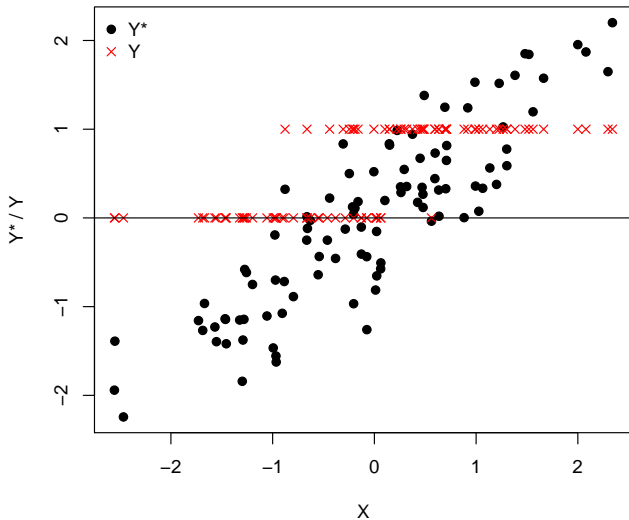
- The Universe: Logit  $>$  Probit
- The (Social Science) Universe: Meh...
- $\hat{\beta}_{\text{Logit}} \approx 1.8 \times \hat{\beta}_{\text{Probit}}$
- Four reasons to prefer / use logit

# A Toy Example

```
> set.seed(7222009)
> ystar<-rnorm(100)
> y<-ifelse(ystar>0,1,0)
> x<-ystar+(0.5*rnorm(100))
> data<-data.frame(ystar,y,x)
> head(data)
```

	ystar	y	x
1	-0.64045247	0	-0.55254581
2	0.58855848	1	1.30215029
3	0.64815988	1	0.70827789
4	-0.50684531	0	0.06377499
5	0.01932982	1	0.63521460

# A Toy Example



# Toy Example: Probit

```
> myprobit<-glm(y~x,family=binomial(link="probit"), data=data)
> summary(myprobit)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.28477	-0.32228	0.00975	0.38602	2.27744

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	0.3228	0.1923	1.679	0.0932 .
x	2.0090	0.3718	5.404	6.51e-08 ***

---

Signif. codes:

0 \*\*\* 0.001 \*\* 0.01 \* 0.05 . 0.1 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 137.989 on 99 degrees of freedom  
Residual deviance: 57.908 on 98 degrees of freedom  
AIC: 61.908

Number of Fisher Scoring iterations: 7

# Toy Example: Logit

```
> mylogit<-glm(y~x,family=binomial(link="logit"), data=data)
> summary(mylogit)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-2.2708	-0.3286	0.0456	0.3934	2.2899

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	0.5320	0.3390	1.569	0.117
x	3.5061	0.7261	4.828	1.38e-06 ***

---

Signif. codes:

0 \*\*\* 0.001 \*\* 0.01 \* 0.05 . 0.1 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 137.989 on 99 degrees of freedom  
Residual deviance: 58.498 on 98 degrees of freedom  
AIC: 62.498

Number of Fisher Scoring iterations: 6

## Toy Example (continued)

Note:

- $zs$ ,  $Ps$ ,  $\ln Ls$  (via “residual deviance”) nearly identical
- $\hat{\beta}_{\text{Logit}}$  is  $\frac{3.5061}{2.0090} = 1.745 \times \hat{\beta}_{\text{Probit}}$

# Toy Example: Predicted Probabilities

