Binary Choice Model:

$$Y_i = \begin{cases} 1 & \text{yes} \\ 0 & \text{no} \end{cases}$$

In the background is some latent variable

$$Y_i^* = X_i'\beta + u_i$$

$$\Rightarrow \quad Y_i = \begin{cases} 1 & Y_i^* > 0 \\ 0 & Y_i^* \leq 0 \end{cases}$$

Wlog, let $c=1$ (this is wlog if X_i includes a constant)

i.e. $Y_i^* = \alpha_0 + X_{1i}'\beta_0 + u_i > c$

iff $\underbrace{(\alpha_0 - c)}_{\alpha_0^*} + X_{1i}'\beta_0 + u_i > 0$

$$\begin{aligned} \Pr[Y_i = 1 | X_i] &= \Pr[Y_i^* > 0 | X_i] \\ &= \Pr[X_i'\beta + u_i > 0 | X_i] \\ &= \Pr[u_i > -X_i'\beta | X_i] \end{aligned}$$

Everything that follows is based on our assumptions about the distribution of u_i

---

LPM

$$E[Y_i | X_i] = X_i'\beta$$

$\Leftrightarrow$ $\boxed{1}$ $Y_i = X_i'\beta + \varepsilon_i$

$\boxed{2}$ $E[\varepsilon_i | X_i] = 0$

How can we ensure that $0 < X_i'\hat{\beta} < 1$?

- We cannot.

Note that $Y_i = 1 \Rightarrow X_i'\beta + \varepsilon_i = 1 \Rightarrow \varepsilon_i = 1 - X_i'\beta$
$Y_i = 0 \Rightarrow X_i'\beta + \varepsilon_i = 0 \Rightarrow \varepsilon_i = -X_i'\beta$

Then $\varepsilon_i | X_i$ has only two support points.

$\Pr[Y_i = 1 | X_i] = X_i'\beta$
$\Pr[Y_i = 0 | X_i] = 1 - X_i'\beta$

$$\Rightarrow \varepsilon_i = \begin{cases} 1 - X_i'\beta & \text{w/prob } X_i'\beta \\ -X_i'\beta & \text{w/prob } 1 - X_i'\beta \end{cases}$$

This is a weird assumption

$V(\varepsilon_i | X_i) = (X_i'\beta)(1 - X_i'\beta)$

$L_n(\beta; \{X_i\}, \{Y_i\}) = \prod_{i=1}^{n} (X_i'\beta)^{Y_i} (1 - X_i'\beta)^{1-Y_i}$

Using MLE, we get $\hat{\beta}$

Logit/Probit: we take as a starting point assumptions about u_i.

$$\begin{aligned} \Pr[Y_i = 1 | X_i] &= \Pr[u_i > -X_i'\beta | X_i] \\ &= 1 - \Pr[u_i \leq -X_i'\beta | X_i] \\ &= 1 - F_{u_i|X_i}(-X_i'\beta) \end{aligned}$$

if $u_i|X_i$ is symmetric around 0. $= F_{u_i|X_i}(X_i'\beta)$

Probit: assume $u_i | X_i \sim N(0, \sigma_u^2)$

$\Rightarrow -u_i | X_i \sim N(0, \sigma_u^2)$

Thus $Pr[Y_i = 1 | X_i] = Pr[-u_i \leq X_i'\beta | X_i]$

$$= Pr\left[\frac{-u_i}{\sigma_u} \leq X_i' \frac{\beta}{\sigma_u} \Big| X_i\right]$$

$$= \Phi\left(X_i' \frac{\beta}{\sigma_u}\right)$$

Here, we can only estimate $\frac{\beta}{\sigma_u}$.

If we assume $u_i | X_i \sim \Lambda$, then $-u_i | X_i \sim \Lambda$

$\Rightarrow Pr[Y_i = 1 | X_i] = Pr[-u_i \leq X_i'\beta | X_i]$

$$= \Lambda(X_i'\beta)$$

$$= \frac{\exp\{X_i'\beta\}}{1 + \exp\{X_i'\beta\}}$$

Recall: Long ago, we estimated

$Y_i = Y_i^* = X_i'\beta + u_i$, $u_i \sim N(0, \sigma_u^2)$

and we estimated $\hat{\beta}$ and $\hat{\sigma}_u^2$.

Why can we identify both here but not in the Probit model? We cannot see Y_i^* is the Probit model, so we cannot estimate σ_u^2

all we see is $Y_i = \tau(Y_i^*)$. Since we do not know τ, we cannot derive $V(Y_i^*)$ from $V(Y_i)$.

$$L_n(\beta \mid \{X_i\},\{Y_i\}) = \prod_{i=1}^{n} [F_{u_i|X_i}(X_i'\beta)]^{Y_i}[1-F_{u_i|X_i}(X_i'\beta)]^{1-Y_i}$$

$$\ln L_n(\beta \mid \{X_i\},\{Y_i\}) = \sum_{i=1}^{n} [Y_i \ln F_{u_i|X_i}(X_i'\beta) + (1-Y_i)\ln F_{u_i|X_i}(X_i'\beta)]$$

The logistic distribution has slightly larger tails.

Semiparametric least squares (late 1980s - early 1990s)

- Hidehiko Ichimura

$\hat{E}[Y_i \mid X_i]$ Nadaraya-Watson

$$\min \sum_{i=1}^{n} (Y_i - \hat{E}[Y_i \mid X_i])^2$$

- $Y_i = X_i'\beta + u_i$ — parametric if we make assumptions about u_i
- $Y_i = X_i'\beta + \varepsilon_i$ — no assumptions about ε_i. semi-parametric
- $Y_i = h(X_i, \theta, \varepsilon_i)$ or $Y_i = h(X_i, \theta) + \varepsilon_i$ — with no assumptions about the function. non-parametric

What can one do with $\hat{\beta}$?

$$\frac{\partial}{\partial X_{ki}} \hat{P}r[Y_i = 1 \mid X_i] = \frac{\partial}{\partial X_{ki}} \Phi(X_i'\hat{\beta}_n) = \varphi(X_i'\hat{\beta})\beta_{nk}$$

This is a function of i. Thus we have to report

$\varphi((\frac{1}{n}\sum_{i=1}^{n} X_i)'\hat{\beta}_n)\hat{\beta}_{nk}$ or $\bar{\eta} = \frac{1}{n}\sum_{i=1}^{n} \varphi(X_i'\hat{\beta}_n)\hat{\beta}_{nk}$ - Buchinsky reports this one!

$$sd(\bar{\eta}) = \frac{1}{n}\sum_{i=1}^{n} (\varphi(X_i'\hat{\beta}_n)\hat{\beta}_{n,k} - \bar{\eta})^2$$