$$L = \sum_{i=1}^{T} y_i \ln\left[F(x_i'\beta)\right] + (1-y_i)\ln\left[1 - F(x_i'\beta)\right]$$

$$\frac{\partial L}{\partial \beta} = \sum_{i=1}^{T} x_i\left[y_i \frac{f_i}{F_i} - (1-y_i)\frac{f_i}{1-F_i}\right]$$

$f_i = f(x_i'\beta)$ pdf

$F_i = F(x_i'\beta)$ cdf

$$\frac{\partial^2 L}{\partial \beta \partial \beta'} = \sum_{i=1}^{T} y_i \left[\frac{F_i \cdot \partial f_i / \partial \beta' - f_i^2}{F_i^2} \right] x_i x_i' - \sum_{i=1}^{T} (1-y_i) \left[\frac{(1-F_i) \cdot \partial f_i / \partial \beta' + f_i^2}{(1-F_i)^2} \right] x_i x_i'$$

$I(\beta) = -E(H)$

AND $E(y_i) = F_i$

$$I(\beta) = \sum_{i=1}^{T} \left[\frac{f_i^2}{(F_i(1-F_i))} \right] x_i x_i'$$

* The first derivative of the log likelihood <u>will not</u> yield an analytical solution for β in terms of y & X.

Use the iterative algorithm

$$\beta_{n+1} = \beta_n - t_n P_n \gamma_n$$

t_n step length

P_N Pos. def matrix evaluated at β_n

γ_n gradient evaluated at β_n

Choose

$$P_n = \left[\frac{\partial^2 \ell}{\partial\beta\,\partial\beta^T}\bigg|_{\beta_n} \right]^{-1} \quad NR$$

$$P_N = -\left[I(\beta) \right]^{-1} \quad MOS$$

$$P_n = \left[\sum_1^{T} \left(\frac{\partial \ell_i}{\partial \beta}\bigg|_{\beta_n} \right) \left(\frac{\partial \ell_i}{\partial \beta}\bigg|_{\beta_n} \right)^T \right]^{-1}$$

BHHH

$\tilde{\beta}$ be MLE

$$\sqrt{n}(\tilde{\beta} - \beta) \xrightarrow{d} N\left(0, \lim_{n\to\infty} \left(\frac{I(\beta)}{n}\right)^{-1}\right)$$

Let $\tilde{\beta}$ be MLE, Then

$$\sqrt{T}\,(\tilde{\beta}-\beta) \xrightarrow{d} N\left(0,\ \lim_{n\to\infty}\left(I(\beta)/n\right)^{-1}\right)$$

For Probit the following results can be used to simplify the Hessian

- $f(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$

- $\frac{\partial f(t)}{\partial t} = -t \cdot f(t)$

$$\Rightarrow \partial f_t / \partial \beta = -x_t' \beta \cdot f(x_t' \beta) \cdot x_t$$

- $F(-t) = 1 - F(t)$

For logit

$$\frac{\partial^2 L}{\partial\beta\,\partial\beta'} = -\sum_{i=1}^{n} f(x_i'\beta)\, x_i x_i'$$

which is the same as

$E(H)$

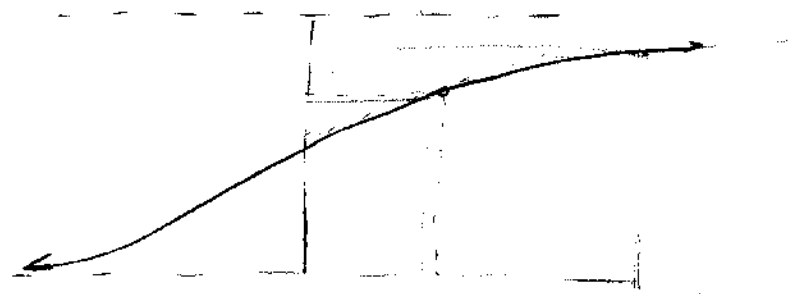
$\therefore$ MOS AND NR are the same for logit!

(1) Interpretation

What does β mean?

In linear model $y = x\beta + \varepsilon$ $\quad \beta_i = \frac{\partial y}{\partial x_i}$

$P_i = F(x_i'\beta)$



$x_i'\beta$

The effect of a change in x on P_i depends on where you are.

Probit

$$\frac{\partial P_i}{\partial x_{ij}} = f(x_i'\beta) \cdot \beta_j$$

: you basically have to choose some representative row / vector and evaluate the effect of Δ in X at that level of X.

(2) Hypoth Tests

$$u = 2 \ln \left[\ell(\hat{\Omega}) - \hat{\ell}(\omega) \right]$$

un $\quad$ rest

$$u = +2 \left[\ln \ell(\tilde{\beta}_{MLE}) - \ln \ell(\hat{\beta}_R) \right] \sim \chi^2_J \text{ if } H_0 \text{ True.}$$

NOTE: Probit $\ell < 0$

pseudo R^2

$$1 - \frac{\ln \hat{\ell}(\Omega) \text{ unrest}}{\ln \hat{\ell}(\omega) \text{ rest}}$$

Also

When $\hat{\beta}_2 = \beta_3 \cdots = \beta_K = 0$
it equals 0 - model has no explan power.

$\doteq 1$ when $\tilde{\beta} = \beta$ $F(x_i\tilde{\beta}) = y_i$
i.e., when the model is perfect predictor.

Note: won't = 1 unless $x'\beta \to \infty$ or $-\infty$
indicating a flawed model
e.g. will happen if dummy var is same as dep. var

TABLE

		Pred $D=0$	Pred $D=1$	
Actual	$D=0$	50	8	58
	$D=1$	6	36	42
		56	44	100

Predicts 86/100 correctly 86%

compare to predict 58% are 0 ⟹ naive 100% '0' 58% right.