Linear probability model: $Pr[Y_i = 1|X_i] = X_i\beta_0$

Logit model: $Pr[Y_i = 1|X_i] = \frac{\exp\{X_i\beta_0\}}{1+\exp\{X_i\beta_0\}}$

Probit model: $Pr[Y_i = 1|X_i] = \Phi(X_i\beta_0)$

<u>Binary choice:</u>

Latent variable interpretation

Probit: $Y_i = 1_{\{X_i\beta_0 - \varepsilon_i > 0\}}$ $\varepsilon_i|X_i \sim N(0,1)$

($X_i\beta_0 - \varepsilon_i$ = Y_i^* ← unobservable)

$\Rightarrow Pr[Y_i = 1|X_i] = \Phi(X_i\beta_0) = E[1\{Y_i = 1\}|X_i]$
$= E[Y_i|X_i]$

Logit: $Y_i = 1_{\{X_i\beta_0 - \varepsilon_i > 0\}}$ $\varepsilon_i|X_i \sim N(0,1)$

$\Rightarrow Pr[Y_i = 1|X_i] = \Lambda(X_i\beta_0)$

We can estimate those parameters using

- MLE
- NLS ⇒ Since the data are heteroskedastic, we have that WNLS is more efficient.
- WNLS

Note that we can only estimate the coefficients "up to scale" meaning that only the ratios and signs have meaning.

<u>Multinomial choice models</u>

$Y_i = 0, 1, 2, \ldots, J_i$ (ie each i has J_i+1 "choices")

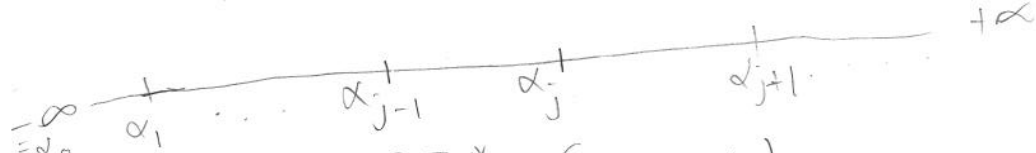
$Pr[Y_i = j|X_i] = F_j(x_i, \theta_0)$, $\forall j = 1, \ldots, J_i$ (← specified as known functions) ($j=0$ is implied as $1 - \sum_{j=1}^{J_i} Pr[Y_i = j|X_i]$)

Define $Y_{ij} = \begin{cases} 1 & \text{if } Y_i = j \\ 0 & \text{else} \end{cases}$

$$\mathcal{L} = \log L = \sum_{i=1}^{n} \sum_{j=0}^{J_i} y_{ij} \ln F_j(x_i, \theta)$$

<u>Ordered Models</u>

$Y_i^* = X_i \beta_0 + \varepsilon_i$ latent variable , $\varepsilon_i | X_i \sim F$



$Y_i = j$ if $Y_i^* \in (\alpha_j, \alpha_{j+1})$

$$\Rightarrow \Pr[Y_i = j | X_i] = \Pr[\alpha_j < Y_i^* < \alpha_{j+1} | X_i]$$
$$= \Pr[\alpha_j < X_i \beta_0 + \varepsilon_i < \alpha_{j+1} | X_i]$$
$$= \Pr[\alpha_j - X_i \beta_0 < \varepsilon_i < \alpha_{j+1} - X_i \beta_0 | X_i]$$
$$= \Pr[\varepsilon_i < \alpha_{j+1} - X_i \beta_0 | X_i] - \Pr[\varepsilon_i < \alpha_j - X_i \beta_0 | X_i]$$
$$= F[\alpha_{j+1} - X_i \beta_0] - F[\underbrace{\alpha_j - X_i \beta_0}_{= (\alpha_j - \beta_{01}) - \tilde{X}_i \tilde{\beta}_0}]$$

The unknowns are the α_j's and β_0.
Can only estimate the difference between the threshold and the constant.

The log likelihood function is globally concave provided f is concave.

The log likelihood fcn here is.

$$\mathcal{L} = \sum_{i=1}^{n} \sum_{j=0}^{J_i} Y_{ij} \ln [F(\alpha_{j+1} - X_i \beta_0) - F(\alpha_j - X_i \beta_0)]$$

Imposing this restriction can cause biases if the true model is unordered.

Unordered Models:

$$U_{ij} = X_i \beta_{0j} + \varepsilon_{ij} \quad , j = 0, 1, \ldots, J_i$$

Choose $Y_{ij} = 1$ if $U_{ij} = \max\{U_{i0}, \ldots, U_{iJ_i}\}$

The most tractable versions of these models are:

Multinomial Logit:

ε_{ij} are iid Type I extreme value: Then,

$$\Pr[Y_i = j \mid X_i] = \frac{\exp\{X_i \beta_{0j}\}}{\sum_{j=0}^{J_i} \exp\{X_i \beta_{0j}\}}$$

- In the Binary case: $\Pr[Y_i = 1 \mid X_i] = \dfrac{\exp\{X_i \beta_0\}}{1 + \exp\{X_i \beta_0\}}$

$$U_{ij} = X_i \beta_{0j} + \varepsilon_{ij}, \quad j = 0, 1$$

$$\Rightarrow \Pr[Y_i = 1 \mid X_i] = \Pr[U_{i1} > U_{i2} \mid X_i]$$

$$= \Pr[X_i \beta_{01} + \varepsilon_{i1} > X_i \beta_{00} + \varepsilon_{i0}]$$

$$= \Pr[X_i(\beta_{01} - \beta_{00}) > \varepsilon_{i0} - \varepsilon_{i1}]$$

$$= \frac{\exp\{X_i(\beta_{01} - \beta_{00})\}}{1 + \exp\{X_i(\beta_{01} - \beta_{00})\}} \left[\frac{\exp\{X_i \beta_{00}\}}{\exp\{X_i \beta_{00}\}}\right]$$

$$= \frac{\exp\{X_i \beta_{01}\}}{\exp\{X_i \beta_{01}\} + \exp\{X_i \beta_{00}\}}$$

Does it make sense to have all ε_{ij} independent?

- This is known as the independence of irrelevant alternatives. (IIA)

That is, $Pr[Y_i = \underbrace{car}_{j} \mid Y_i = \underbrace{car}_{j} \text{ or } \underbrace{bus}_{k}] = \dfrac{\exp\{X_i\beta_{0j}\}}{\exp\{X_i\beta_{0j}\} + \exp\{X_i\beta_{0k}\}}$

McFadden generalizes this model to overcome this assumption by assuming

$$\varepsilon_{ij} \sim \text{Type B extreme value}$$

Then they are correlated. This is known as the 2-level Nested Logit Model.

$$F(\varepsilon_{i1}, \varepsilon_{i2}) = \exp\{-[\exp\{-\rho^{-1}\varepsilon_{i1}\} + \exp\{-\rho^{-1}\varepsilon_{i2}\}]^{\rho}\}$$

where $corr(\varepsilon_{i1}, \varepsilon_{i2}) = 1 - \rho^2$. This is very difficult to estimate.

Multivariate Probit model.

$$\begin{bmatrix} \varepsilon_{i0} \\ \vdots \\ \varepsilon_{iJ_i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \Sigma \right)$$

eg: $J_i = 2$

$$Pr[Y_i = 2 \mid X_i] = Pr[U_{i2} > U_{i1}, U_{i2} > U_{i0}]$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{U_{i1}} \int_{-\infty}^{U_{i0}} f(v_1, v_2, v_3 \mid X_i)\, dv_0\, dv_1\, dv_2$$