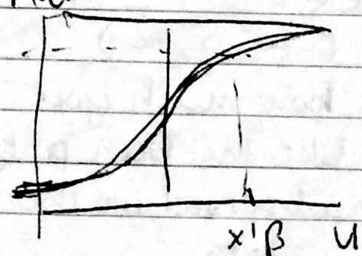
$Pr(y=1|x)$
$\wr$
$F(u)$

$Pr(y=1|x) = F(x'\beta)$



Random Utility Model

$y = 1$ if alternative 1 provides higher utility

0 if alt 0 provides higher utility

Additive rand utility model

$U_0 = V_0 + \varepsilon_0$

$U_1 = \underbrace{V_1}_{\text{deterministic comp of utility}} + \underbrace{\varepsilon_1}_{\text{random comp of utility}}$

If we observe $y = 1 \Rightarrow U_1 > U_0$

$$\begin{aligned} Pr(y=1) &= Pr(U_1 > U_0) \\ &= Pr(V_1 + \varepsilon_1 > V_0 + \varepsilon_0) \\ &= Pr(\varepsilon_0 - \varepsilon_1 < V_1 - V_0) \\ &= F(V_1 - V_0) \end{aligned}$$

where F is cdf of $(\varepsilon_0 - \varepsilon_1)$

$= F(x'\beta)$

if $V_1 - V_0 = x'\beta$

Ordered multinomial

$y_i^* = x_i'\beta + u_i$

as y_i^* cross threshold we move up the ordering of alternatives

Ex. y^* : continuous happiness
y : survey answer

$y_i = j$ if $\alpha_{j-1} < y_i^* < \alpha_j$

what's random?

$$\begin{aligned} Pr[y_i = j] &= Pr(\alpha_{j-1} < y_i^* < \alpha_j) \\ &= Pr(\alpha_{j-1} < x_i'\beta + u_i < \alpha_j) \\ &= Pr(\alpha_{j-1} - x_i'\beta < u_i < \alpha_j - x_i'\beta) \\ &= \cancel{Pr}\,F(\alpha_j - x_i'\beta) - F(\alpha_{j-1} - x_i'\beta) \end{aligned}$$

where $F(\cdot)$ is the cdf of u_i

How many unknowns? For m choice, have $\underbrace{k}_{\text{regressor}\ \beta} + \underbrace{m-1}_{\text{cutpoint}}$

How to estimate this w/ ML?

$$L_N = \prod_{\underbrace{i=1}_{\text{individual}}}^{N} \prod_{\underbrace{j=1}_{\text{choices}}}^{m} p_{ij}^{y_{ij}}$$

$p_{ij} \leftarrow$ what's unknown?
y_{ij} what's observed?

$$\mathcal{L} = \ln L_N = \sum\sum y_{ij} \ln \underbrace{p_{ij}}_{=F_j(x_i,\beta)}$$

different models use different $F_j(x_i,\beta)$

$$\frac{\partial L}{\partial \beta} = \sum\sum \frac{y_{ij}}{p_{ij}} \cdot \frac{\partial p_{ij}}{\partial \beta} = 0$$

Ordered logit $u \sim \text{logit}$ $F(z) = \frac{e^z}{1+e^z}$

probit $u \sim \Phi$

Interpretation

Sign of β = whether y^* increases w/ the regressor

$$\frac{\partial \Pr(y_i = j)}{\partial x_i} = \{F'(\alpha_{j-1} - x_i'\beta) - F'(\alpha_j - x_i'\beta)\}\beta$$

w^2

$\frac{p}{1-p}$

Ordered logit:

$$\Pr(y_i = j) = F(\alpha_j - x_i'\beta) - F(\alpha_{j-1} - x_i'\beta)$$
$$= \frac{\exp(\alpha_j - x_i'\beta)}{1+\exp(\alpha_j - x_i'\beta)} - \frac{\exp(\alpha_{j-1} - x_i'\beta)}{1+\exp(\alpha_{j-1} - x_i'\beta)}$$

if $j = 1$ vs 0 (logit is a special case of ordered logit)

$$\Pr(y_i = j) = \frac{\exp(\infty)}{1+\exp(\infty)} - \frac{\exp(-x_i'\beta)}{1+\exp(\cancel{\alpha_j} - x_i'\beta)}$$
$$= 1 - \frac{\exp(-x_i'\beta)}{1+\exp(-x_i'\beta)} = \frac{1}{1+\exp(-x_i'\beta)}$$

✓

~~logit $p(\cancel{\mp})$ $\frac{\exp(x_i'\beta)}{1+\exp(x_i'\beta)}$~~

Proportional odd assumption

poor, fair, good, very good, excellent

p_1 p_2 p_3 p_4 p_5

logodd of

$$\text{poor} = \log \frac{p_1}{p_2 + p_3 + \cancel{p_4 + p_5}}$$

$$\text{poor or fair} = \log \frac{p_1 + p_2}{p_3 + \cancel{p_4 + p_5}}$$

$$= \frac{p_1 + p_2 + p_3}{p_4 + p_5}$$

same

a linear combo of x

$$\frac{p_1}{p_2 + p_3 + \cancel{p_4 + p_5}} = \frac{F(\alpha_1 - x_i'\beta) - F(\alpha_0 - x_i'\beta)}{F(\alpha_2 - x_i'\beta) - F(\alpha_1 - x_i'\beta) + F(\alpha_3 - x_i'\beta) - F(\alpha_2 - x_i'\beta)}$$

$$\log \frac{p_1 + p_2}{p_3} - \log \frac{p_1}{p_2 + p_3}$$

$$= \log\left(\frac{p_1 + p_2}{p_3} \cdot \frac{p_2 + p_3}{p_1}\right)$$

$$\log(1) - \log\left(\frac{p_1 + p_2}{p_3}\right)$$

$$= \log\left(\frac{p_3}{p_1 + p_2}\right)$$

$$\log \frac{p_1+p_2}{p_3} - \log \frac{p_1}{p_2+p_3}$$
$$= \alpha_2 - x_i'\beta - \log(1+\alpha_2 - x_i'\beta)$$

$$\log(p_1+p_2) = \log\big(F(\alpha_1 - x_i'\beta) - F(\cancel{\alpha}\, x_i'\beta) + F(\alpha_2 - \cancel{x}_i'\beta) - F(\alpha_1 - x_i'\beta)\big)$$
$$= \log\big(F(\alpha_2 - x_i'\beta)\big)$$

$$\log(p_2+p_3) = \log\big(\cancel{F}(\alpha_3 - x_i'\beta) - F(\alpha_1 - x_i'\beta)\big)$$

$$\log \frac{p_1+p_2}{p_3} = \log\big(F(\alpha_2 - x_i'\beta)\big) - \log\big(F(\alpha_3 - x_i'\beta) - F(\alpha_2 - x_i'\beta)\big)$$

$$\log \frac{p_1}{p_2+p_3} = \log\big(F(\alpha_1 - x_i'\beta)\big) - \log\big(F(\alpha_3 - x_i'\beta) - F(\alpha_1 - x_i'\beta)\big)$$

$$\log(1) = 0$$

$$\log \frac{p_1+p_2}{p_3} - \log \frac{p_1}{p_2+p_3} =$$

Recall $F(z) = \frac{e^z}{1+e^z}$

$$\log F(z) = z - \log(1+e^z)$$

Prop odds

$$\log\left(\frac{\Pr(Y \le k)}{\Pr(Y > k)}\right) = \alpha_k - x'\beta \quad \forall k$$

$\beta > 0 \Rightarrow x\uparrow$ ~~means~~ $Y\uparrow$

Simplify β - only 1 set

$$\Pr(Y \le k) = \Pr(\mu_i \le \alpha_k - x_i'\beta)$$
$$= F(\alpha_j - x_i'\beta)$$
$$= \frac{\exp(\alpha_j - x\beta)}{1 + \exp(\alpha_j - x\beta)}$$

$$\Pr(Y > k) = \Pr(\mu_i > \alpha_k - x_i'\beta)$$
$$= 1 - \Pr(\mu_i < \alpha_k - x_i'\beta)$$
$$= 1 - F(\alpha_k - x_i'\beta)$$
$$= \frac{1}{1 + \exp(\alpha_j - x\beta)}$$

$$\log\left(\frac{P(Y \le k)}{P(Y > k)}\right) = \alpha_k - x_i'\beta$$