Binary choice model

Logit & Probit model

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

$$g(z) = \frac{e^z}{(1+e^z)^2}$$

$$\Phi(z) \equiv \int_{-\infty}^{z} \phi(v)dv$$

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

Coefficient estimates and standard errors in the logit model are roughly $\pi/\sqrt{3} = 1.8$ bigger than in the probit model.

Odds ratio= $\frac{P(y=1|x=1)}{P(y=1|x=0)}$

Latent variable

y_i^* is unobserved, it is referred to as a latent variable.

$$y_i^* = x_i'\beta + \varepsilon_i. \qquad (7.8)$$

$$\begin{aligned} y_i &= 1 \quad \text{if } y_i^* > 0 \\ &= 0 \quad \text{if } y_i^* \leq 0, \end{aligned} \qquad (7.10)$$

$$P\{y_i = 1\} = P\{y_i^* > 0\} = P\{x_i'\beta + \varepsilon_i > 0\} = P\{-\varepsilon_i \leq x_i'\beta\} = F(x_i'\beta), \qquad (7.9)$$

Estimation:

$$L(\beta) = \prod_{i=1}^{N} P\{y_i = 1|x_i;\beta\}^{y_i} P\{y_i = 0|x_i;\beta\}^{1-y_i}, \qquad (7.11)$$

$$\log L(\beta) = \sum_{i=1}^{N} y_i \log F(x_i'\beta) + \sum_{i=1}^{N} (1-y_i)\log(1-F(x_i'\beta)). \qquad (7.12)$$

$$\frac{\partial \log L(\beta)}{\partial \beta} = \sum_{i=1}^{N} \left[\frac{y_i - F(x_i'\beta)}{F(x_i'\beta)(1-F(x_i'\beta))} f(x_i'\beta) \right] x_i = 0, \qquad (7.13)$$

Goodness-of-fit:
Goodness-of-fit measures are based on comparison with a model that contains only a constant as explanatory variable. Let logL1 denote the maximum loglikelihood value of the model of interest and let logL0 denote the maximum value of the loglikelihood function when all parameters, except the intercept, are set to zero. Clearly, logL1 ≥

logL0. The larger the difference between the two loglikelihood values, the more the extended model adds to the very restrictive model.

1.
$$pseudoR^2 = 1 - \frac{1}{1 + 2(\log L_1 - \log L_0)/N}, \tag{7.17}$$

2.
$$McFaddenR^2 = 1 - \log L_1 / \log L_0, \tag{7.18}$$

Table 7.1 Cross-tabulation of actual and predicted outcomes

		$\hat{y}_i$		
		0	1	Total
y_i	0	n_{00}	n_{01}	N_0
	1	n_{10}	n_{11}	N_1
	Total	n_0	n_1	N

$$\hat{p} = N_1/N,$$

the proportion of incorrect predictions:

$$wr_1 = \frac{n_{01} + n_{10}}{N},$$

$$\begin{aligned} wr_0 &= 1 - \hat{p} \quad \text{if } \hat{p} > 0.5, \\ &= \hat{p} \quad\quad\; \text{if } \hat{p} \le 0.5. \end{aligned}$$

3.

$$R_p^2 = 1 - \frac{wr_1}{wr_0}. \tag{7.21}$$

Illustration: the Impact of Unemployment Benefits on Recipiency

Table 7.2 Binary choice models for applying for unemployment benefits (blue collar workers)

	LPM		Logit		Probit	
Variable	Estimate	s.e.	Estimate	s.e.	Estimate	s.e.
constant	−0.077	(0.122)	−2.800	(0.604)	−1.700	(0.363)
replacement rate	0.629	(0.384)	3.068	(1.868)	1.863	(1.127)
replacement rate^2	−1.019	(0.481)	−4.891	(2.334)	−2.980	(1.411)
age	0.0157	(0.0047)	0.068	(0.024)	0.042	(0.014)
$age^2/10$	−0.0015	(0.0006)	−0.0060	(0.0030)	−0.0038	(0.0018)
tenure	0.0057	(0.0012)	0.0312	(0.0066)	0.0177	(0.0038)
slack work	0.128	(0.014)	0.625	(0.071)	0.375	(0.042)
abolished position	−0.0065	(0.0248)	−0.0362	(0.1178)	−0.0223	(0.0718)
seasonal work	0.058	(0.036)	0.271	(0.171)	0.161	(0.104)
head of household	−0.044	(0.017)	−0.211	(0.081)	−0.125	(0.049)
married	0.049	(0.016)	0.242	(0.079)	0.145	(0.048)
children	−0.031	(0.017)	−0.158	(0.086)	−0.097	(0.052)
young children	0.043	(0.020)	0.206	(0.097)	0.124	(0.059)
live in SMSA	−0.035	(0.014)	−0.170	(0.070)	−0.100	(0.042)
non-white	0.017	(0.019)	0.074	(0.093)	0.052	(0.056)
year of displacement	−0.013	(0.008)	−0.064	(0.015)	−0.038	(0.009)
>12 years of school	−0.014	(0.016)	−0.065	(0.082)	−0.042	(0.050)
male	−0.036	(0.018)	−0.180	(0.088)	−0.107	(0.053)
state max. benefits	0.0012	(0.0002)	0.0060	(0.0010)	0.0036	(0.0006)
state unempl. rate	0.018	(0.003)	0.096	(0.016)	0.057	(0.009)
Loglikelihood			−2873.197		−2874.071	
Pseudo R^2			0.066		0.066	
McFadden R^2			0.057		0.057	
R_p^2	0.035		0.046		0.045	

The replacement rate, defined as the ratio of weekly UI(unemployment insurance) benefits to previous weekly earnings.

The LPM is estimated by OLS, the logit and probit model are both estimated by maximum likelihood. The estimates of β obtained from the logit model are roughly a factor $\pi/\sqrt{3}$ larger than those obtained from the probit model.

The replacement rate has an insignificant positive coefficient, while its square is significantly negative. For the probit model, we can derive that the estimated marginal effect of a change in the replacement rate (rr) equals **the value of the normal density function** multiplied by $1.863 - 2 \times 2.980rr$.

The dummy variable which indicates whether the job was lost because of slack work is highly significant in all specifications, which is not surprising given that these workers typically will find it hard to get a new job.

The higher the state unemployment rate and the higher the maximum benefit level, the more likely it is that individuals apply for benefits.

The ceteris paribus effect of being married is estimated to be positive, while, somewhat surprisingly, being head of the household has a negative effect on the probability of take-up.

Table 7.3 Cross-tabulation of actual and predicted outcomes (logit model)

		$\hat{y}_i$		
		0	1	Total
y_i	0	242	1300	1542
	1	171	3164	3335
	Total	413	4464	4877

$$R_p^2 = 1 - \frac{171 + 1300}{1542},$$

$\hat{p} = 3335/4877$

$$\log L_0 = 3335 \log \frac{3335}{4877} + 1542 \log \frac{1542}{4877} = -3046.187,$$

which allows us to compute the pseudo and McFadden R2 measures.

$p_{00} + p_{11} =$

$$\frac{242}{1542} + \frac{3164}{3335} = 1.106,$$

Multi-response Models

Ordered Response Models

$$y_i^* = x_i'\beta + \varepsilon_i \tag{7.28}$$

$$y_i = j \quad \text{if } \gamma_{j-1} < y_i^* \le \gamma_j, \tag{7.29}$$

$$y_i^* = x_i'\beta + \varepsilon_i \tag{7.30}$$

$$\begin{aligned} y_i &= 1 \quad \text{if } y_i^* \le 0, \\ &= 2 \quad \text{if } 0 < y_i^* \le \gamma, \\ &= 3 \quad \text{if } y_i^* > \gamma, \end{aligned} \tag{7.31}$$

Assuming that εi is i.i.d. standard normal results in the ordered probit model. The logistic distribution gives the ordered logit model. For M = 2 we are back at the binary choice model.

$$P\{y_i = 1|x_i\} = P\{y_i^* \le 0|x_i\} = \Phi(-x_i'\beta),$$
$$P\{y_i = 3|x_i\} = P\{y_i^* > \gamma|x_i\} = 1 - \Phi(\gamma - x_i'\beta)$$

and

$$P\{y_i = 2|x_i\} = \Phi(\gamma - x_i'\beta) - \Phi(-x_i'\beta),$$

Normalization:

$$P\{y_i = 1|x_i\} = P\{\beta_1 + x_i'\beta + \varepsilon_i \le \gamma_1|x_i\} = \Phi\left(\frac{\gamma_1 - \beta_1}{\sigma} - x_i'\left(\frac{\beta}{\sigma}\right)\right),$$

Illustration: Willingness to Pay for Natural Areas

Table 7.4 Ordered probit model for willingness to pay

	I: intercept only		II: with characteristics	
Variable	Estimate	s.e.	Estimate	s.e.
constant	18.74	(2.77)	30.55	(8.59)
age class	–		−6.93	(1.64)
female	–		−5.88	(5.07)
income class	–		4.86	(1.87)
$\hat{\sigma}$	38.61	(2.11)	36.47	(1.89)
Loglikelihood	−409.00		−391.25	
Normality test (χ_2^2)	6.326	($p = 0.042$)	2.419	($p = 0.298$)

Multinomial Models

$$\begin{aligned} P\{y_i = j\} &= P\{U_{ij} = \max\{U_{i1}, \dots, U_{iM}\}\} \\ &= P\left\{\mu_{ij} + \varepsilon_{ij} > \max_{k=1,\dots,J,k\neq j}\{\mu_{ik} + \varepsilon_{ik}\}\right\}. \end{aligned} \qquad (7.35)$$

we assume that all εij are mutually independent with a so-called log Weibull distribution (also known as a Type I extreme value distribution).

$$F(t) = \exp\{-e^{-t}\}, \qquad (7.36)$$

$$P\{y_i = j\} = \frac{\exp\{\mu_{ij}\}}{\exp\{\mu_{i1}\} + \exp\{\mu_{i2}\} + \cdots + \exp\{\mu_{iM}\}}. \qquad (7.37)$$

multinomial logit model:

$$P\{y_i = j\} = \frac{\exp\{x_{ij}'\beta\}}{1 + \exp\{x_{i2}'\beta\} + \cdots + \exp\{x_{iM}'\beta\}}, \quad j = 1, 2, \dots, M. \qquad (7.38)$$

independence of irrelevant alternatives (IIA)

The probability ratio (or odds ratio) is given by:

$$\frac{P(y_i = j)}{P(y_i = k)} = \frac{exp(x_{ij}'\beta)}{exp(x_{ik}'\beta)}$$

Models for Count Data

Poisson regression model

$$P\{y_i = y|x_i\} = \frac{\exp\{-\lambda_i\}\lambda_i^y}{y!}, \quad y = 0, 1, 2, \ldots, \tag{7.42}$$

estimation:

$$\log L(\beta) = \sum_{i=1}^{N} [-\lambda_i + y_i \log \lambda_i - \log y_i!] \tag{7.44}$$
$$= \sum_{i=1}^{N} [-\exp\{x_i'\beta\} + y_i(x_i'\beta) - \log y_i!].$$

equidispersion:

$$V\{y_i|x_i\} = \exp\{x_i'\beta\}. \tag{7.43}$$

overdispersion:

$$V\{y_i|x_i\} = E\{\varepsilon_i^2|x_i\} > \exp\{x_i'\beta\}$$

NegBin I model:

$$V\{y_i|x_i\} = (1 + \delta^2)\exp\{x_i'\beta\} \tag{7.49}$$

NegBin II model:

$$V\{y_i|x_i\} = (1 + \alpha^2 \exp\{x_i'\beta\})\exp\{x_i'\beta\}, \tag{7.50}$$

Illustration: Patents and R&D Expenditures

Table 7.5 Estimation results Poisson model, MLE and QMLE

	MLE		QMLE
	Estimate	Standard error	Robust s.e.
constant	−0.8737	0.0659	0.7429
log ($R\&D$)	0.8545	0.0084	0.0937
aerospace	−1.4218	0.0956	0.3802
chemistry	0.6363	0.0255	0.2254
computers	0.5953	0.0233	0.3008
machines	0.6890	0.0383	0.4147
vehicles	−1.5297	0.0419	0.2807
Japan	0.2222	0.0275	0.3528
USA	−0.2995	0.0253	0.2736
Loglikelihood	−4950.789		
Pseudo R^2	0.675		
LR test (χ^2_8)	20587.54 ($p = 0.000$)		
Wald test (χ^2_8)			338.9 ($p = 0.000$)

Table 7.6 Estimation results NegBin I and NegBin II model, MLE

	NegBin I (MLE)			NegBin II (MLE)	
	Estimate	Standard error		Estimate	Standard error
constant	0.6899	0.5069		−0.3246	0.4982
log ($R\&D$)	0.5784	0.0676		0.8315	0.0766
aerospace	−0.7865	0.3368		−1.4975	0.3772
chemistry	0.7333	0.1852		0.4886	0.2568
computers	0.1450	0.2063		−0.1736	0.2988
machines	0.1559	0.2550		0.0593	0.2793
vehicles	−0.8176	0.2686		−1.5306	0.3739
Japan	0.4005	0.2573		0.2522	0.4264
USA	0.1588	0.1984		−0.5905	0.2788
δ^2	95.2437	14.0069	α^2	1.3009	0.1375
Loglikelihood	−848.195			−819.596	
Pseudo R^2	0.944			0.946	
LR test (χ^2_8)	88.55 ($p = 0.000$)			145.75 ($p = 0.000$)	