Chap 6. Multicategory Logit Models

Y has J categories, $J > 2$.
Sampling model : At each setting or predictors, we assume indep. multinomial distributions with probabilities.

Model for Nominal Response

Let

$$\pi_j = P(Y = j),\ j = 1,\cdots,J$$

Baseline-category logit model

$$log\left(\frac{\pi_j}{\pi_J}\right) = \alpha_j + \beta_j x,\ j = 1,2,\cdots,J-1\ .$$

i.e., Separate set of parameters (α_j, β_j) occurs for each logit (for each predictor)

Note:

- Which category we use as the baseline category (i.e., category J) is arbitrary (For nominal variables, the order of the categories is arbitrary).
- $\exp(\hat{\beta}_j)$ is the multiplicative impact of a 1-unit increase in x on the odds of making response j instead of response J.
- Can use model with ordinal response variables also, but then you ignore information about ordering

Ex) Income and Job satisfaction (1991, GSS data)

	Job satisfaction Very Dissat	Little Dissat	Moderate Satisfied	Very Satisfied	Total
< 5,000	2	4	13	3	22
5,000~15,000	2	6	22	4	34
15,000~25,000	0	1	15	8	24
> 25,000	0	3	13	8	24
Total	4	14	63	23	104

Using $x =$ income scores (3,10,20,30), we use SAS (PROC CATMOD or LOGISTIC) to fit model:

Prediction equations

$$\log\frac{\hat{\pi}_1}{\hat{\pi}_4} = 0.56 - 0.20x$$

$$\log\frac{\hat{\pi}_2}{\hat{\pi}_4} = 0.65 - 0.07x$$

$$\log\frac{\hat{\pi}_3}{\hat{\pi}_4} = 1.82 - 0.05x$$

Note:

- For each logit, odds of being in less satisfied category (instead of very satisfied) decrease as $x = income \uparrow$
- The estimated odds of being "very dissatisfied" instead of "very satisfied" multiplies by $e^{-0.20} = 0.82$ for each 1 thousand dollar increase in income.

  For a 10 thousand dollar increase in income; (eg, from row 2 to row 3 or from row 3 to row 4 of table), the estimated odds multiply by $e^{10(-0.20)} = e^{-2.0} = 0.14$.

  eg.) at $x = 30$, the estimated odds of being "very dissatisfied" instead of "very satisfied" are just 0.14 times the corresponding odds at $x = 20$.
- Model treats income as quantitative, $Y =$ job satisfaction as qualitative (nominal), but Y is ordinal (We later consider a model that treats job satisfaction as ordinal)

Estimating response probabilities.

Equivalent form of model is

$$\pi_j = \frac{e^{\alpha_j + \beta_j x}}{1 + e^{\alpha_1 + \beta_1 x} + \cdots + e^{\alpha_{J-1} + \beta_{J-1} x}}, \ j = 1, 2, \cdots, J-1$$

Then

$$\frac{\pi_j}{\pi_J} = e^{\alpha_j + \beta_j x}$$

Note $\sum_{j=1}^{N} \pi_j = 1$.

Ex) Job satisfaction data

$$\hat{\pi}_1 = \frac{e^{0.56 - 0.20x}}{1 + e^{0.56 - 0.20x} + e^{0.65 - 0.07x} + e^{0.1.82 - 0.05x}}$$

$$\hat{\pi}_2 = \frac{e^{0.65 - 0.07x}}{1 + e^{0.56 - 0.20x} + e^{0.65 - 0.07x} + e^{0.1.82 - 0.05x}}$$

$$\hat{\pi}_3 = \frac{e^{0.1.82 - 0.05x}}{1 + e^{0.56 - 0.20x} + e^{0.65 - 0.07x} + e^{0.1.82 - 0.05x}}$$

$$\hat{\pi}_4 = \frac{1}{1 + e^{0.56 - 0.20x} + e^{0.65 - 0.07x} + e^{0.1.82 - 0.05x}}$$

eg) at $x=30$, estimated porb. of “very satisfied” is

$$\hat{\pi}_4 = \frac{1}{1+e^{0.56-0.20(30)}+e^{0.65-0.07(30)}+e^{0.1.82-0.05(30)}} = 0.365$$

Likewise, $\hat{\pi}_1 = 0.002$, $\hat{\pi}_2 = 0.084$, $\hat{\pi}_3 = 0.550$.

$$\hat{\pi}_1 + \hat{\pi}_2 + \hat{\pi}_3 + \hat{\pi}_4 = 1.0$$

● ML estimates determine effects for all pairs of categories,

$$\begin{aligned} \text{eg) } \log\left(\frac{\hat{\pi}_1}{\hat{\pi}_2}\right) &= \log\left(\frac{\hat{\pi}_1}{\hat{\pi}_4}\right) - \log\left(\frac{\hat{\pi}_1}{\hat{\pi}_4}\right) \\ &= (0.564-0.199x)-(0.645-0.070x) \\ &= -0.081-0.129x \end{aligned}$$

● Since we have a contingency table data, we can test goodness of fit
The Deviance is the LR test statistic for testing data all parameters not in model=0.

$Deviance = G^2 = 4.18$, $df=6$, $p-value = 0.65$ for H_0 : Model hols with linear trends for income. (Also, Pearson $X^2 = 3.6$, $df = 6$, $p-value = 0.73$ for same hypothesis)
Model has 12 logits (3 at each of income levels), 6 parameters, so $df = 12-6=6$ for testing fit.

Note: Inference uses usual methods

● Wald C.I. for β_j is $\hat{\beta}_j \pm Z_{\alpha/2}(SE)$

● Wald test of $H_0 : \beta_j = 0$ uses $Z = \frac{\hat{\beta}_j}{SE}$ or $Z^2 \sim \chi_1^2$

● For small n, better to use LR test and LR C.I.

ex) Overall “global” test of income effect

$$H_0 : \beta_1 = \beta_2 = \beta_3 = 0$$

SAS reports Wald stat.=7.03, df=3, p-vale=0.07 (see model ❶)
Weak evidence, but ignores ordering of satisfaction categories (With many parameters, Wald test.=quandratic form $\underline{\hat{\beta}}^T[cov(\underline{\hat{\beta}})]^{-1}\underline{\hat{\beta}}$)
Can get LR statistic by comparing Deviance with simpler “independence model”

$LR\ stat. = 9.29,\ df = 3,\ p - vale = 0.03$
(9.29 = 13.47(❸) - 4.18(❶)

Model for Ordinal Response

The cumulative probabilities are

$$P(Y \le j) = \pi_1 + \cdots + \pi_j,\ \ j = 1, 2, \cdots, J$$

The cumulative link model are

$$G^{-1}(P(Y \le j)) = \alpha_j + \beta x$$

1) Cumulative logit model: G^{-1} is the logit function;
2) Cumulative probit model: $G = \Phi$ (cdf of $N(0,1)$);
3) Cumulative complementary log-log link: G=extreme value cdf.

Cumulative logits are

$$\begin{aligned} logit P(Y \le j) &= \log\left[\frac{P(Y \le j)}{1 - P(Y \le j)}\right] = \log\left[\frac{P(Y \le j)}{P(Y > j)}\right] \\ &= \log\left[\frac{\pi_1 + \cdots + \pi_j}{\pi_{j+1} + \cdots + \pi_J}\right] \end{aligned}$$

for $j = 1, 2, \cdots, J-1$

Cumulative logit model has form

$$logit P(Y \le j) = \alpha_j + \beta x$$

- Seperate intercept α_j for each cumulative logit.
- same slope β for each cumulative logit

Note:

- e^{β} = multiplicative effect of 1-unit change in x on odds that $P(Y \le j)$ (instead of $P(Y > j)$)
- $\dfrac{odds\ \ of\ \ P(Y \le j)\ \ at\ \ x_2}{odds\ \ of\ \ P(Y \le j)\ \ at\ \ x_1} = e^{\beta(x_2 - x_1)}$

  Also called Proportional odds model
- SAS: ML fit with PROC LOGISTIC or PROC GENMOD(dist=mult, link=clogit), PROC LOGISTIC default for dummy variable is 1 in category, -1 if in last category, 0 otherwise.
  To use usual form of 1 in category, 0 oterwise, use param=ref option. eg.
  CLASS race gender/param=ref;

```
data jobsatis;
 input income satis count @@;
cards;
3  1 2  3 2 4  3 3 13  3 4 3
10 1 2 10 2 6 10 3 22 10 4 4
20 1 0 20 2 1 20 3 15 20 4 8
30 1 0 30 2 3 30 3 13 30 4 8
;
/* Baseline - Category logit */
❶ proc catmod data=jobsatis;
    weight count;
  /* The  'direct'  statement treats the indep. variable(income) as quantitative variable*/
    response logits direct income;
    model satis=income/pred=freq;
   run;

/* Cumulative logit */
❷ proc genmod data=jobsatis;
    weight count;
    model satis=income/dist=mult link=clogit type3;
   run;

/* Indepence as baseline-category logit model */
❸ proc catmod data=jobsatis;
    weight count;
  /* The  'population' statement specifies that populations are to be based only on
    cross-classifications of the specified variables */
    response logits population income;
    model satis=/pred=freq;
   run;

/* Independence as loglinear model */
❹ proc genmod data=jobsatis;
    class income satis;
    model count=income satis /dist=poi link=log;
   run;
```

The CATMOD Procedure

Data Summary

Response	satis	Response Levels	4
Weight Variable	count	Populations	4
Data Set	JOBSATIS	Total Frequency	104
Frequency Missing	0	Observations	14

Population Profiles

Sample	income	Sample Size
1	3	22
2	10	34
3	20	24
4	30	24

Response Profiles

Response	satis
1	1
2	2
3	3
4	4

Maximum Likelihood Analysis

Maximum likelihood computations converged.

Maximum Likelihood Analysis of Variance

Source	DF	Chi-Square	Pr > ChiSq
Intercept	3	15.30	0.0016
income	3	7.03	0.0709
Likelihood Ratio	6	4.18	0.6528

Analysis of Maximum Likelihood Estimates

Parameter	Function Number	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	1	0.5638	0.9601	0.34	0.5570
	2	0.6451	0.6688	0.93	0.3347
	3	1.8187	0.5290	11.82	0.0006
income	1	-0.1988	0.1021	3.79	0.0515
	2	-0.0705	0.0370	3.64	0.0564
	3	-0.0469	0.0255	3.38	0.0660

The CATMOD Procedure

Maximum Likelihood Predicted Values for Response Functions

		------Observed------		------Predicted-----		
income	Function Number	Function	Standard Error	Function	Standard Error	Residual
3	1	-0.40547	0.912871	-0.0325	0.749942	-0.37297
	2	0.287682	0.763763	0.433585	0.578126	-0.1459
	3	1.466337	0.640513	1.677943	0.462793	-0.21161
10	1	-0.69315	0.866025	-1.42391	0.673296	0.73076
	2	0.405465	0.645497	-0.05993	0.407545	0.465395
	3	1.704748	0.543557	1.349515	0.326586	0.355233
20	1	.	.	-3.41164	1.438541	.
	2	-2.07944	1.06066	-0.76495	0.39755	-1.31449
	3	0.628609	0.437798	0.880333	0.252516	-0.25172
30	1	.	.	-5.39937	2.402098	.
	2	-0.98083	0.677003	-1.46997	0.650472	0.489144
	3	0.485508	0.449359	0.41115	0.388732	0.074357

Maximum Likelihood Predicted Values for Frequencies

income	satis	Frequency	Observed Standard Error	Predicted Frequency	Predicted Standard Error	Residual
3	1	2	1.3484	2.402231	1.365717	-0.40223
	2	4	1.809068	3.82852	1.378488	0.17148
	3	13	2.306118	13.28767	1.777047	-0.28767
	4	3	1.60963	2.481575	0.998831	0.518425
10	1	2	1.371989	1.355749	0.803155	0.644251
	2	6	2.222876	5.303311	1.380771	0.696689
	3	22	2.786522	21.71008	1.883047	0.289916
	4	4	1.878673	5.630856	1.500012	-1.63086
20	1	0	0	0.202476	0.286428	-0.20248
	2	1	0.978945	2.856374	0.899768	-1.85637
	3	15	2.371708	14.80311	1.286808	0.196886
	4	8	2.309401	6.138035	1.125808	1.861965
30	1	0	0	0.039543	0.09414	-0.03954
	2	3	1.620185	2.011795	1.10503	0.988205
	3	13	2.44097	13.19913	2.130443	-0.19913
	4	8	2.309401	8.749534	2.106893	-0.74953

The GENMOD Procedu

Model Information

Data Set	WORK.JOBSATIS
Distribution	Multinomial
Link Function	Cumulative Logit
Dependent Variable	satis
Scale Weight Variable	count
Number of Observations Read	16
Number of Observations Used	14
Sum of Weights	104
Missing Values	2

Response Profile

Ordered Value	satis	Total Frequency	Total Weight
1	1	2	4
2	2	4	14
3	3	4	63
4	4	4	23

PROC GENMOD is modeling the probabilities of levels of satis having LOWER Ordered Values in the response profile table.
One way to change this to model the probabilities of HIGHER Ordered Values is to specify the DESCENDING option in the PROC statement.

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Log Likelihood		-103.6335	
Full Log Likelihood		-103.6335	
AIC (smaller is better)		215.2670	
AICC (smaller is better)		219.7115	
BIC (smaller is better)		217.8233	

Algorithm converged.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept1	1	-2.4732	0.5713	-3.5930	-1.3534	18.74	<.0001
Intercept2	1	-0.7817	0.3776	-1.5218	-0.0416	4.29	0.0384
Intercept3	1	2.2111	0.4466	1.3358	3.0864	24.51	<.0001
income	1	-0.0563	0.0210	-0.0976	-0.0151	7.17	0.0074
Scale	0	1.0000	0.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

The GENMOD Procedure

LR Statistics For Type 3 Analysis

Source	DF	Chi-Square	Pr > ChiSq
income	1	7.51	0.0061

The CATMOD Procedure

Data Summary

Response	satis	Response Levels	4
Weight Variable	count	Populations	4
Data Set	JOBSATIS	Total Frequency	104
Frequency Missing	0	Observations	14

Population Profiles

Sample	income	Sample Size
1	3	22
2	10	34
3	20	24
4	30	24

Response Profiles

Response	satis
1	1
2	2
3	3
4	4

Maximum Likelihood Analysis
Maximum likelihood computations converged.

Maximum Likelihood Analysis of Variance

Source	DF	Chi-Square	Pr > ChiSq
Intercept	3	56.06	<.0001
Likelihood Ratio	9	13.47	0.1426

Analysis of Maximum Likelihood Estimates

Parameter	Function Number	Estimate	Standard Error	Chi-Square	Pr > ChiSq
Intercept	1	-1.7492	0.5417	10.43	0.0012
	2	-0.4964	0.3390	2.14	0.1431
	3	1.0076	0.2436	17.11	<.0001

The CATMOD Procedure

Maximum Likelihood Predicted Values for Response Functions

income	Function Number	------Observed------ Function	Standard Error	------Predicted----- Function	Standard Error	Residual
3	1	-0.40547	0.912871	-1.7492	0.541733	1.343735
	2	0.287682	0.763763	-0.49644	0.338979	0.784119
	3	1.466337	0.640513	1.007641	0.243621	0.458697
10	1	-0.69315	0.866025	-1.7492	0.541733	1.056053
	2	0.405465	0.645497	-0.49644	0.338979	0.901902
	3	1.704748	0.543557	1.007641	0.243621	0.697108
20	1	.	.	-1.7492	0.541733	.
	2	-2.07944	1.06066	-0.49644	0.338979	-1.583
	3	0.628609	0.437798	1.007641	0.243621	-0.37903
30	1	.	.	-1.7492	0.541733	.
	2	-0.98083	0.677003	-0.49644	0.338979	-0.48439
	3	0.485508	0.449359	1.007641	0.243621	-0.52213

Maximum Likelihood Predicted Values for Frequencies

income	satis	-------Observed------ Frequency	Standard Error	------Predicted------ Frequency	Standard Error	Residual
3	1	2	1.3484	0.846154	0.414859	1.153846
	2	4	1.809068	2.961538	0.736305	1.038462
	3	13	2.306118	13.32692	1.054229	-0.32692
	4	3	1.60963	4.865385	0.895322	-1.86538
10	1	2	1.371989	1.307692	0.641145	0.692308
	2	6	2.222876	4.576923	1.137927	1.423077
	3	22	2.786522	20.59615	1.629262	1.403846
	4	4	1.878673	7.519231	1.383679	-3.51923
20	1	0	0	0.923077	0.452573	-0.92308
	2	1	0.978945	3.230769	0.803242	-2.23077
	3	15	2.371708	14.53846	1.150068	0.461538
	4	8	2.309401	5.307692	0.976715	2.692308
30	1	0	0	0.923077	0.452573	-0.92308
	2	3	1.620185	3.230769	0.803242	-0.23077
	3	13	2.44097	14.53846	1.150068	-1.53846
	4	8	2.309401	5.307692	0.976715	2.692308

The GENMOD Procedure

Model Information

Data Set	WORK.JOBSATIS
Distribution	Poisson
Link Function	Log
Dependent Variable	count

Number of Observations Read	16
Number of Observations Used	16

Class Level Information

Class	Levels	Values
income	4	3 10 20 30
satis	4	1 2 3 4

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	9	13.4673	1.4964
Scaled Deviance	9	13.4673	1.4964

Pearson Chi-Square	9	11.5242	1.2805
Scaled Pearson X2	9	11.5242	1.2805
Log Likelihood		129.0550	
Full Log Likelihood		-31.5342	
AIC (smaller is better)		77.0684	
AICC (smaller is better)		91.0684	
BIC (smaller is better)		82.4765	

Algorithm converged.

Analysis Of Maximum Likelihood Parameter Estimates

Parameter		DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept		1	1.6692	0.2748	1.1305	2.2078	36.89	<.0001
income	3	1	-0.0870	0.2952	-0.6655	0.4915	0.09	0.7682
income	10	1	0.3483	0.2666	-0.1742	0.8708	1.71	0.1914
income	20	1	0.0000	0.2887	-0.5658	0.5658	0.00	1.0000
income	30	0	0.0000	0.0000	0.0000	0.0000	.	.
satis	1	1	-1.7492	0.5417	-2.8110	-0.6874	10.43	0.0012
satis	2	1	-0.4964	0.3390	-1.1608	0.1679	2.14	0.1431
satis	3	1	1.0076	0.2436	0.5302	1.4851	17.11	<.0001
satis	4	0	0.0000	0.0000	0.0000	0.0000	.	.
Scale		0	1.0000	0.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

Ex) Job statisfaction and income (See model ❷ in SAS output)

$$\text{logit}\hat{P}(Y \le j) = \hat{\alpha}_j + \hat{\beta}x = \hat{\alpha}_j - 0.056x,\ j = 1,2,3$$

Odds of response at low end of job satisfaction scale ↓ as $x = income \uparrow$

$$e^{\hat{\beta}} = e^{-0.056} = 0.95$$

Estimated odds of satisfaction below any given level (instead of above it) multiplies by 0.95 for 1-unit increase in x (but $x = 3,10,20,30$).

For $10,000 increase in income, estimated odds multiply by

$$e^{10\hat{\beta}} = e^{10(-0.056)} = 0.57$$

eg. estimated odds of satisfaction being below (instead of above) some level at $30,000 income equal 0.57 times the odds at $20,000.

Note:

● If reverse order, $\hat{\beta}$ change sign but has same SE

ex) Category 1=Very satisfied, 2=Moderately satisfied, 3=Little dissatisfied, 4=Very dissatisfied

$$\hat{\beta} = 0.056,\ e^{\hat{\beta}} = 1.06 = 1/0.95$$

(Response mode likely at "very satisfied" end of scale as $x \uparrow$)

● $H_0 : \beta = 0$ (job satisfaction indep. of income) has

$$Wald\ stat. = \left(\frac{\hat{\beta} - 0}{SE}\right)^2 = \left(\frac{-0.056}{0.021}\right)^2 = 7.17,\ df = 1,\ p-value = 0.007$$

$LR\ stat. = 7.51\,(df = 1,\ p-vale = 0.006)$

(LR stat. from Type3 option)

These tests give stronger evidence of association than if treat:

● Y as nominal (BCL model)

$$\log\frac{\pi_j}{\pi_4} = \alpha_j + \beta_j x$$

(Recall $p-value = 0.07$ for Wald test of $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$)

● X, Y as nominal

Pearson test of indep. has $X^2 = 11.5$, $df = 9$, $p-value = 0.24$ (analogous to testing all $\beta_j = 0$ in BCL model with dummy predictors)

With BCL or cumulative logit models, we can have quantitative and qualitative predictors, interaction terms, etc.

Ex) GSS data

$Y=$ political ideology (1=Very liberal, ... , 5=very conservative)

x_1 = gender (1=F, 0=M)

x_2 = political party (1=Democrat, 0=Republican)

Gender	Politcal Party	Political ideology				
		Very liberal	Slightly liberal	Moderate	Slightly conservative	Very conservative
Female	Democratic	44	47	118	23	32
	Republican	18	28	86	39	48
Male	Democratic	36	34	53	18	23
	Republican	12	18	62	45	51

● ML fit

$$logit\,\hat{P}(Y \le j) = \hat{\alpha}_j + 0.117x_1 + 0.964x_2$$

For $\hat{\beta}_1 = 0.117$, $SE = 0.127$

For $\hat{\beta}_2 = 0.964$, $SE = 0.130$

For each gender, estimated odds for a Democrat's response is in liberal rather than conservative direction (i.e., $Y \le j$ rather than $Y > j$) are $e^{0.964} = 2.62$ times estimated odds for Republican's response.

● 95% C.I. for true odds ratio is

$$e^{0.964 \pm 1.96(0.130)} = (2.0,\ 3.4)$$

● LR test of $H_0 : \beta_2 = 0$ (no party effect, given gender) has

$$test\ \ stat. = 56.8,\ df = 1\,(p-value < 0.0001)$$

⇒ Vey strong evidence that Democrats tend to be more liberal than Republicans(for each gender).

● Not much evidence of gender effect (for each party) But, is there interaction?

ML fit of model permitting interaction is

$$logit\,\hat{P}(Y \le j) = \hat{\alpha}_j + 0.366x_1 + 1.265x_2 - 0.509x_1x_2$$

Estimated odds ratio for party effect (x_2) is

$$e^{1.265} = 3.5 \ \ when \ \ x_1 = 0(M)$$
$$e^{1.265-0.509} = 2.2 \ \ when \ \ x_1 = 1(F)$$

Estimated odds ratio for gender effect (x_1) is

$e^{0.366} = 1.4$ when $x_2 = 0$ (Republican)

$e^{0.366-0.509} = 0.9$ when $x_2 = 1$ (Democrat)

i.e., for Republican ($x_2 = 0$), females ($x_1 = 1$) tend to be more liberal than male ($x_1 = 0$)

$$logit\hat{P}(Y \le j) = \hat{\alpha}_j + 0.366$$

Find $\hat{P}(Y=1)$ (very liberal) for male Republicans, female Republicans

$$\hat{P}(Y \le j) = \frac{e^{\hat{\alpha}_j + 0.366x_1 + 1.265x_2 - 0.509x_1x_2}}{1 + e^{\hat{\alpha}_j + 0.366x_1 + 1.265x_2 - 0.509x_1x_2}}$$

For $j=1$, $\hat{\alpha}_1 = -2.674$

$$\hat{P}(Y=1) = \frac{e^{-2.674}}{1+e^{-2.674}} = 0.064 \ \ for \, male \ \ Republican$$
$$\hat{P}(Y=1) = \frac{e^{-2.674+0.366}}{1+e^{-2.674+0.366}} = 0.090 \ \ for \, female \ \ Republican$$

(Weak gender effect for Republicans, likewise for Democrats but in opposite direction)

Similarly, $\hat{P}(Y=2) = \hat{P}(Y \le 2) - \hat{P}(Y \le 1), \ etc.$,

$\hat{P}(Y=5) = 1 - \hat{P}(Y \le 4)$

Note:

- If reverse order of response categories, estimates change sign and odds ratio → 1/(odds ratio)
  (very liberal, ... , very conservative ⇒ very conservative, ... , very liberal)
- For ordinal response, other orders are not sensible.
  ex) categories (liberal, moderate, conservative)
  Enter into SAS as 1,2,3
  or PROC GENMOD ORDER=DATA
  or else SAS will alphabetize as (conservative, lineral, moderate) and treat that as ordering for the cumulative logits.

GSS_politics

```
data ideology;
input gender party ideology count @@;
datalines;
1 1 1 44   1 1 2 47   1 1 3 118   1 1 4 23   1 1 5 32
1 2 1 18   1 2 2 28   1 2 3  86   1 2 4 39   1 2 5 48
2 1 1 36   2 1 2 34   2 1 3  53   2 1 4 18   2 1 5 23
2 2 1 12   2 2 2 18   2 2 3  62   2 2 4 45   2 2 5 51
;

/* Cumulative logit model with gender and party */
❶-1 proc logistic;
       class gender(ref=last) party(ref=last)/param=ref;
       weight count;
       model ideology = gender party;
     run;

❶-2 proc genmod;
       class gender(ref=last) party(ref=last)/param=ref;
       weight count;
       model ideology = gender party / dist=mult link=clogit type3;
     run;

/* Cumulative logit model with gender, party, and interaction */
❷-1 proc logistic;
       class gender(ref=last) party(ref=last)/param=ref;
       weight count;
       model ideology = gender party gender*party;
     run;

❷-2 proc genmod;
       class gender(ref=last) party(ref=last)/param=ref;
       weight count;
       model ideology = gender party gender*party/ dist=mult link=clogit type3;
     run;
```

$$gender = \begin{cases} 1 & Female \\ 2 & Male \end{cases}$$

$$party = \begin{cases} 1 & Democrate \\ 2 & Republican \end{cases}$$

The LOGISTIC Procedure

Model Information

Data Set	WORK.IDEOLOGY
Response Variable	ideology
Number of Response Levels	5
Weight Variable	count
Model	cumulative logit
Optimization Technique	Fisher's scoring

Number of Observations Read	20
Number of Observations Used	20
Sum of Weights Read	835
Sum of Weights Used	835

Response Profile

Ordered Value	ideology	Total Frequency	Total Weight
1	1	4	110.00000
2	2	4	127.00000
3	3	4	319.00000
4	4	4	125.00000
5	5	4	154.00000

Probabilities modeled are cumulated over the lower Ordered Values.

Class Level Information

Class	Value	Design Variables
gender	1	1
	2	0
party	1	1
	2	0

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

The LOGISTIC Procedure

Score Test for the Proportional Odds Assumption

Chi-Square	DF	Pr > ChiSq
11.0066	6	0.0882

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	2541.630	2486.142
SC	2545.613	2492.117
-2 Log L	2533.630	2474.142

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	59.4878	2	<.0001
Score	58.2776	2	<.0001
Wald	57.5831	2	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
gender	1	0.8498	0.3566
party	1	55.4882	<.0001

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1	-2.5323	0.1495	286.7605	<.0001
Intercept	2	1	-1.5388	0.1295	141.1003	<.0001
Intercept	3	1	0.1745	0.1166	2.2381	0.1346
Intercept	4	1	1.0085	0.1243	65.8614	<.0001
gender	1	1	0.1169	0.1268	0.8498	0.3566
party	1	1	0.9636	0.1294	55.4882	<.0001

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
gender 1 vs 2	1.124	0.877	1.441
party 1 vs 2	2.621	2.034	3.377

Association of Predicted Probabilities and Observed Responses

Percent Concordant	37.5	Somers' D	0.000
Percent Discordant	37.5	Gamma	0.000
Percent Tied	25.0	Tau-a	0.000
Pairs	160	c	0.500

The GENMOD Procedure

Model Information

Data Set	WORK.IDEOLOGY
Distribution	Multinomial
Link Function	Cumulative Logit
Dependent Variable	ideology
Scale Weight Variable	count

Number of Observations Read	20
Number of Observations Used	20
Sum of Weights	835

Class Level Information

Class	Value	Design Variables
gender	1	1
	2	0
party	1	1
	2	0

Response Profile

Ordered Value	ideology	Total Frequency
1	1	110
2	2	127
3	3	319
4	4	125
5	5	154

PROC GENMOD is modeling the probabilities of levels of ideology having LOWER Ordered Values inthe response profile table. One way to change this to model the probabilities of HIGHER OrderedValues is to specify the DESCENDING option in the PROC statement.

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Log Likelihood		-1237.0711	

The GENMOD Procedure
Algorithm converged.
Analysis Of Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept1	1	-2.5322	0.1489	-2.8242	-2.2403	289.05	<.0001
Intercept2	1	-1.5388	0.1297	-1.7931	-1.2845	140.67	<.0001
Intercept3	1	0.1745	0.1162	-0.0533	0.4023	2.25	0.1332
Intercept4	1	1.0086	0.1232	0.7672	1.2499	67.07	<.0001
gender 1	1	0.1169	0.1273	-0.1327	0.3664	0.84	0.3588
party 1	1	0.9636	0.1297	0.7095	1.2178	55.22	<.0001
Scale	0	1.0000	0.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

LR Statistics For Type 3 Analysis

Source	DF	Chi-Square	Pr > ChiSq

gender	1	0.84	0.3586
party	1	56.85	<.0001

The LOGISTIC Procedure

Model Information

Data Set	WORK.IDEOLOGY
Response Variable	ideology
Number of Response Levels	5
Weight Variable	count
Model	cumulative logit
Optimization Technique	Fisher's scoring

Number of Observations Read	20
Number of Observations Used	20
Sum of Weights Read	835
Sum of Weights Used	835

Response Profile

Ordered Value	ideology	Total Frequency	Total Weight
1	1	4	110.00000
2	2	4	127.00000
3	3	4	319.00000
4	4	4	125.00000
5	5	4	154.00000

Probabilities modeled are cumulated over the lower Ordered Values.

Class Level Information

Class	Value	Design Variables
gender	1	1
	2	0
party	1	1
	2	0

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

The LOGISTIC Procedure

Score Test for the Proportional Odds Assumption

Chi-Square	DF	Pr > ChiSq
11.3986	9	0.2494

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	2541.630	2484.150
SC	2545.613	2491.120
-2 Log L	2533.630	2470.150

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	63.4800	3	<.0001
Score	61.4897	3	<.0001
Wald	61.8399	3	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
gender	1	4.1495	0.0416
party	1	41.1818	<.0001
gender*party	1	4.0111	0.0452

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1	-2.6743	0.1660	259.5564	<.0001
Intercept	2	1	-1.6772	0.1482	128.0493	<.0001
Intercept	3	1	0.0424	0.1352	0.0982	0.7541
Intercept	4	1	0.8789	0.1405	39.1252	<.0001
gender	1	1	0.3660	0.1797	4.1495	0.0416
party	1	1	1.2651	0.1971	41.1818	<.0001
gender*party1	1	1	-0.5089	0.2541	4.0111	0.0452

The LOGISTIC Procedure

Association of Predicted Probabilities and Observed Responses

Percent Concordant	37.5	Somers' D	0.000
Percent Discordant	37.5	Gamma	0.000
Percent Tied	25.0	Tau-a	0.000
Pairs	160	c	0.500

The GENMOD Procedure

Model Information

Data Set	WORK.IDEOLOGY
Distribution	Multinomial
Link Function	Cumulative Logit
Dependent Variable	ideology
Scale Weight Variable	count

Number of Observations Read	20
Number of Observations Used	20
Sum of Weights	835

Class Level Information

Class	Value	Design Variables
gender	1	1
	2	0
party	1	1
	2	0

Response Profile

Ordered Value	ideology	Total Frequency
1	1	110
2	2	127
3	3	319
4	4	125
5	5	154

PROC GENMOD is modeling the probabilities of levels of ideology having LOWER Ordered Values inthe response profile table. One way to change this to model the probabilities of HIGHER OrderedValues is to specify the DESCENDING option in the PROC statement.

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Log Likelihood		-1235.0750	

The GENMOD Procedure

Algorithm converged.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept1	1	-2.6743	0.1655	-2.9987	-2.3500	261.20	<.0001
Intercept2	1	-1.6772	0.1476	-1.9665	-1.3880	129.17	<.0001
Intercept3	1	0.0424	0.1338	-0.2198	0.3046	0.10	0.7513
Intercept4	1	0.8790	0.1389	0.6068	1.1512	40.06	<.0001
gender 1	1	0.3661	0.1784	0.0164	0.7157	4.21	0.0402
party 1	1	1.2653	0.1995	0.8743	1.6564	40.21	<.0001
gender*party1 1	1	-0.5091	0.2550	-1.0090	-0.0093	3.99	0.0459
Scale	0	1.0000	0.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

LR Statistics For Type 3 Analysis

Source	DF	Chi-Square	Pr > ChiSq
gender	1	4.22	0.0400
party	1	40.84	<.0001
gender*party	1	3.99	0.0457

Note that c**umulative probit model**s are similar to cumulative logit model.

<u>Complementary log-log link model</u>:

$$\log\{-\log(1-P(Y \le j))\} = \alpha_j + \beta x$$

satisfies

$$P(Y > j | x_1) = \{P(Y > j | x_2)\}^{\exp(\beta(x_2 - x_1))}$$

Ex.) Life length

Y=Life time; $G=1$ (male), 0 (female); $R=1$ (black), 0 (white).

Model:

$$\log\{-\log(1-P(Y \le j))\} = \alpha_j + \beta_1 G + \beta_2 R$$

Life length distribution, in Percentages, of US Residents in 1981

Life length	Males		Females	
	White	Black	White	Black
0-20	2.4 (2.4)	3.6 (4.4)	1.6 (1.2)	2.7 (2.3)
20-40	3.4 (3.5)	7.5 (6.4)	1.4 (1.9)	2.9 (3.4)
40-50	3.8 (4.4)	8.3 (7.7)	2.2 (2.4)	4.4 (4.3)
50-60	17.5 (16.7)	25.0 (26.1)	9.9 (9.6)	16.3 (16.3)
Over 60	72.9 (73.0)	55.6 (55.4)	84.9 (84.9)	73.7 (73.7)

Values in parentheses give fit of proportional hazards model.

Estimated values for β_1 and β_2 are

$$\hat{\beta}_1 = .658 \ (e^{\hat{\beta}_1} = 1.93),$$

$$\hat{\beta}_2 = .626 \ (e^{\hat{\beta}_2} = 1.87).$$

For a given race,

$$\hat{P}(Y > j | male) = \hat{P}(Y > j | female)^{1.93}.$$

Given gender,

$$\hat{P}(Y > j | black) = \hat{P}(Y > j | white)^{1.87}.$$

```
data lifetime;
  input life gender race count @@;
  cards;
  1 1 0 2.4    1 1 1 3.6    1 0 0 1.6    1 0 1 2.7
  2 1 0 3.4    2 1 1 7.5    2 0 0 1.4    2 0 1 2.9
  3 1 0 3.8    3 1 1 8.3    3 0 0 2.2    3 0 1 4.4
  4 1 0 17.5  4 1 1 25.0   4 0 0 9.9    4 0 1 16.3
  5 1 0 72.9  5 1 1 55.6   5 0 0 84.9 5 0 1 73.7
run;

proc logistic data=lifetime;
  weight count;
  model life=gender race / link=cloglog;
run;
```

The LOGISTIC Procedure

Model Information

Data Set	WORK.LIFETIME
Response Variable	life
Number of Response Levels	5
Weight Variable	count
Model	cumulative cloglog
Optimization Technique	Fisher's scoring

Number of Observations Read	20
Number of Observations Used	20
Sum of Weights Read	400
Sum of Weights Used	400

Response Profile

Ordered Value	life	Total Frequency	Total Weight
1	1	4	10.30000
2	2	4	15.20000
3	3	4	18.70000
4	4	4	68.70000
5	5	4	287.10000

Probabilities modeled are cumulated over the lower Ordered Values.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Score Test for the Equal Slopes Assumption

Chi-Square	DF	Pr > ChiSq
0.9994	6	0.9856

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	729.833	711.717
SC	733.816	717.691
-2 Log L	721.833	699.717

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	D	Pr > ChiSq
Likelihood Ratio	22.1162	2	<.0001
Score	21.5473	2	<.0001
Wald	21.2947	2	<.0001

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1	-4.3881	0.3609	147.8372	<.0001
Intercept	2	1	-3.4576	0.2683	166.0191	<.0001
Intercept	3	1	-2.8760	0.2345	150.4058	<.0001
Intercept	4	1	-1.8122	0.1989	83.0321	<.0001
gender		1	0.6577	0.1957	11.2991	0.0008
race		1	0.6264	0.1950	10.3216	0.0013

Association of Predicted Probabilities and Observed Responses

Percent Concordant	37.5	Somers' D	0.000

Percent Discordant	37.5	Gamma	0.000
Percent Tied	25.0	Tau-a	0.000
Pairs	160	c	0.500