Session 6
Proportional odds model
(model checking and interpretation)

6.1 Observed and expected frequencies

6.2 Grouped data

6.3 Ungrouped data

6.4 Stratified proportional odds model

MRC Hubs for Trials Methodology Research
Edinburgh Hub

MRC Hubs for Trials Methodology Research
North West Hub

6.1 Observed and expected frequencies

PROC LOGISTIC will create an output data set

```
output out=fitted xbeta=xbeta stdxbeta=stdxbeta
       predprobs=individual;
```

- (m –1) rows created for each original row of data

- data set includes

(1) **original data**

(2) **xbeta** : estimate of the linear predictor

$$\alpha_1 + \eta(\underline{z}_1)$$

$$\vdots$$

$$\vdots$$

$$\alpha_{m-1} + \eta(\underline{z}_1)$$

$$\Updownarrow$$

$$\alpha_1 + \eta(\underline{z}_n)$$

$$\vdots$$

$$\vdots$$

$$\alpha_{m-1} + \eta(\underline{z}_n)$$

$$\eta(\underline{z}_i) = \beta_1 z_{i1} + \ldots + \beta_q z_{iq}$$

(3) **stdxbeta** : standard error of **xbeta**

(4) If option **predprobs = cumulative**

$$\hat{Q}_k(\underline{z}_i), \quad k = 1,\dots, m$$

where

$$\hat{Q}_k(\underline{z}_i), \text{ is fitted value of } Q_k(\underline{z}_i) = P(C_k \text{ or better; } \underline{z}_i)$$

$$\hat{Q}_k(\underline{z}_i) = (1 + e^{-(\hat{\alpha}_k + \hat{\eta}(\underline{z}_i))})^{-1}$$

$$= (1 + e^{-\text{xbeta}})^{-1}$$

$$Q_m(\underline{z}_i) = 1$$

(5) If option **predprobs = individual**

$$\hat{p}_k(\underline{z}_i), \quad k = 1,..., m$$

where $\hat{p}_k(\underline{z}_i)$ is the predicted probability that outcome is category k for subject i

$$\hat{p}_k(\underline{z}_i) = \hat{Q}_k(\underline{z}_i) - \hat{Q}_{k-1}(\underline{z}_i)$$

6.2 Grouped data

Head injury data from Example 1

GOS at 3 months	GCS on entry		Total
	3-5	6-8	
1. Good Recovery	73	219	292
2. Moderate Disability	55	118	173
3. Severe Disability	79	66	145
4. Vegetative State	37	10	47
5. Dead	358	92	450
Total	602	505	1107

Fitted probabilities

Using output data set

- Fitted probabilities can be multiplied by frequency of respective GCS group to give **expected** frequencies

- Produce table of **observed** and **expected** frequencies

Glasgow Outcome Scale (GOS) at 3 months	Glasgow Coma Scale (GCS) on entry			
	3-5		6-8	
	Obs	Exp	Obs	Exp
1. Good Recovery	73	64.39	219	224.36
2. Moderate Disability	55	67.63	118	104.99
3. Severe Disability	79	81.22	66	67.31
4. Vegetative State	37	30.38	10	17.14
5. Dead	358	358.40	92	91.20
Total	602	602	505	505

Observed and Expected frequencies are not very close

Score test for proportional odds

From SAS output

Score Test for the Proportional Odds Assumption

Chi-Square	DF	Pr > ChiSq
11.8845	3	0.0078

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
-2 Log L	3117.079	2850.319

Our fitted model was:

$$Q_k(\underline{z}_i), \quad = \left[1 + e^{-[\alpha_k + \eta(\underline{z}_i)]}\right]^{-1}, \quad k = 1, \ldots, m-1$$

where $\eta(\underline{z}_i) = \beta z_i$ $\qquad z_i = \begin{cases} 0: \text{ if gcsentry} = 1 \text{ (GCS 3-5)} \\ 1: \text{ if gcsentry} = 2 \text{ (GCS 6-8)} \end{cases}$

$\Rightarrow$ 5 parameters:

$\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4$: intercepts

β : gcsentry

$$-2\text{Log}\hat{L}_{\text{PROP}} = 2850.319$$

Alternative generalised cumulative logit model:

$$Q_k(\underline{z}_i), \quad = \left[1 + e^{-[\alpha_k + \eta_k(\underline{z}_i)]}\right]^{-1}, \quad k = 1, \ldots, m-1$$

where $\eta_k(\underline{z}_i) = \beta_k z_i$

$\Rightarrow$ 8 parameters:

$\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4$: intercepts

$\beta_1 \quad \beta_2 \quad \beta_3 \quad \beta_4$: gcsentry

For the generalised cumulative logit model fitted probabilities are:

Glasgow Outcome Scale (GOS) at 3 months	Glasgow Coma Scale (GCS) on entry	
	3-5	6-8
1. Good Recovery	0.121	0.434
2. Moderate Disability	0.091	0.234
3. Severe Disability	0.131	0.131
4. Vegetative State	0.061	0.020
5. Dead	0.595	0.182

$$-2\log \hat{L}_{GEN} = -2[73\log 0.121 + 219\log 0.434 + \ldots + 92\log 0.182]$$

$$= 2840.068$$

Likelihood ratio test for additional 3 parameters

$$= -2\text{Log}\hat{L}_{\text{PROP}} - -2\text{Log}\hat{L}_{\text{GEN}}$$

= 2850.319 – 2840.068

= 10.251 $(\text{cf. } \chi^2_3, \quad p = 0.0165)$

↗

Score test for proportional odds approximates this as

11.88 $(\text{cf. } \chi^2_3, \quad p = 0.0078)$

Notes:

(1) In this case the generalised cumulative logit model is the saturated model

(2) The generalised cumulative logit model cannot be fitted in SAS PROC LOGISTIC. It is **not** the model fitted by Model option GLOGIT

(3) Test is *very sensitive* when samples are large

6.3 Ungrouped data

Head injury Example 2

Glasgow Outcome Scale	Treatment		Total
	Control	Treated	
1: Good recovery	42	71	113
2: Moderate disability	27	30	57
3: Severe disability	33	27	60
4: Vegetative state/Dead	63	48	111
Total	165	176	341

Covariates: age (yrs) continuous

gcsmotor: 1 = None or extension

2 = Abnormal or normal flexion

3 = Localises

treat: 0 = Control

1 = Treated

patient	age	gcsmotor	treat	gos4
1	66	1	0	1
2	63	1	1	1
3	58	1	1	1
4	54	1	0	1
5	51	1	1	1
6	49	1	0	1
7	49	1	1	1
8	46	1	0	1
9	46	1	1	1
10	43	1	0	1

etc.

Fitting model

$$\log\left[\frac{Q_k(\underline{z}_i)}{1-Q_k(\underline{z}_i)}\right] = \alpha_k + \eta(\underline{z}_i), \quad k = 1,\ldots,m-1$$

$$\eta(\underline{z}_i) = \beta_1 z_{i1} + \beta_2 z_{i2} + \beta_3 z_{i3} + \beta_4 z_{i4}$$

where

$$z_{i1} = \text{ age}$$

$$z_{i2} = \begin{cases} 1: \text{ if gcsmotor} = 2 \text{ (Flexion)} \\ 0: \text{ otherwise} \end{cases} \qquad z_{i3} = \begin{cases} 1: \text{ if gcsmotor} = 3 \text{ (Localises)} \\ 0: \text{ otherwise} \end{cases}$$

$$z_{i4} = \begin{cases} 1: \text{ if treat} = 1 \text{ (Treated)} \\ 0: \text{ if treat} = 0 \text{ (Control)} \end{cases}$$

For this Proportional odds model a SAS Proc Logistic program and output are shown in Supplement 6.3 and 6.3.1

To assess the fit of the model, compare each patient's ***actual*** category with the k ***fitted*** probabilities.

- Group data by the fitted linear predictor: $\eta(\underline{z}_i)$

  Group G_s is all patients with $\eta(\underline{z}_i) \in (a_{s-1}, a_s)$

i.e.

Group G_s	$\eta(\underline{z}_i)$
1	> 0.65
2	> 0.15 to ≤ 0.65
3	> -0.35 to ≤ 0.15
4	> -0.85 to ≤ -0.35
5	≤ -0.85

(Ashby, Pocock and Shaper, 1986; similar to Hosmer and Lemeshow, 1980)

- Count ***observed*** number of patients in G_s with category C_k. This is O_{ks}

- For ***expected*** number of patients in G_s with category C_k, for each category sum the probabilities $\hat{p}_k(\underline{z}_i)$ over all patients in G_s . This is E_{ks}

- Tabulate and compare O_{ks} and E_{ks} *(see Supplement 6.3.1)*

- Could calculate $\chi^2 = \sum (O_{ks} - E_{ks})^2 / E_{ks}$
  (Ashby, Pocock and Shaper, 1986)

- Identify and interpret discrepancies

Score test for proportional odds

Score Test for the Proportional Odds Assumption

Chi-Square	DF	Pr > ChiSq
5.4898	8	0.7042

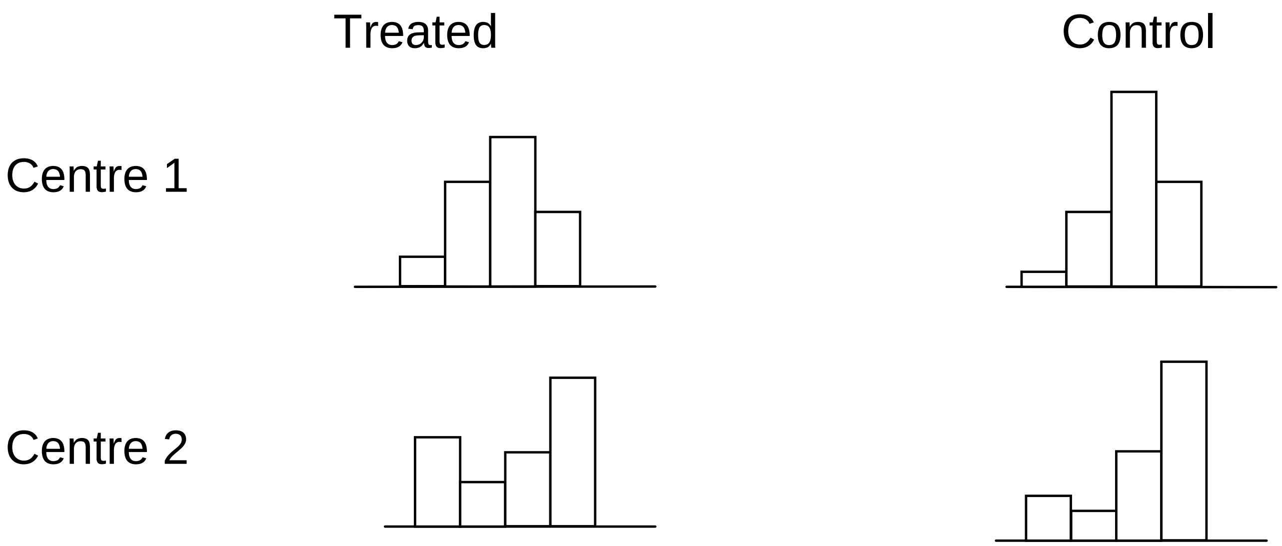
$$\chi^2 = \sum \left(O_{ks} - E_{ks}\right)^2 \big/ E_{ks} = 7.731 \qquad \text{c.f. } \chi^2_8, \ p = 0.46$$

This test is unnecessary given the score test for proportional odds

6.4 Stratified proportional odds model

Odds proportional between treatments, **not** between other prognostic factors

e.g. 2 treatments, 2 centres



So fit stratified POM, stratifying for centre

Stratified proportional odds model:

$$\log\left(\frac{Q_{sk}(z_i)}{1-Q_{sk}(z_i)}\right) = \alpha_{sk} + \theta z_{si}$$

for m categories with k dichotomisations $k = 1,\ldots, m-1$
subject i within stratum s $i = 1,\ldots, n_s$
with treatment indicator variable z_{si} $s = 1,\ldots, r$
where the strata are combinations of prognostic variables

For each dichotomisation of the categories:

- parameter estimate θ is constant
- parameter estimate for each stratum is different

e.g. $\alpha_{11}, \alpha_{12}, \alpha_{13}$: centre 1
$\alpha_{21}, \alpha_{22}, \alpha_{23}$: centre 2 } strata
θ : treatment

- For s^{th} stratum calculate Z_s and V_s (Session 3)

$$Z_s = \frac{1}{n_s + 1} \sum_{k=1}^{m} n_{skC} (L_{skT} - U_{skT})$$

$$V_s = \frac{n_{sT} n_{sC} n_s}{3(n_s + 1)^2} \left(1 - \sum_{k=1}^{m} \left(\frac{n_{sk}}{n_s} \right)^3 \right)$$

Approximately $Z_s \sim N(\theta V_s, V_s)$

- Calculate $Z = \sum_{s=1}^{r} Z_s, \quad V = \sum_{s=1}^{r} V_s$

Then $Z \sim N(\theta V, V)$

Estimation

If $\theta \approx 0 \Rightarrow \frac{Z}{V}$ is an estimate of θ

$$\mathrm{se}\left(\frac{Z}{V}\right) = \frac{1}{\sqrt{V}}$$

Approximate 95% confidence interval for θ:

$$\left[\frac{Z}{V} \pm 1.96\left(\frac{1}{\sqrt{V}}\right)\right]$$

Hypothesis test

To test $\theta = 0$ $\frac{Z^2}{V} \sim \chi_1^2$

This is a meta-analysis (See Session 12)

To fit stratified POM using SAS PROC NLMIXED

(related to partial proportional odds model, Peterson and Harrell, 1990)

Example: Clinical Global Impression of Change Scale (CGIC) in Alzheimer's disease (*Whitehead et al, (2001))*

Response: Clinical Global Impression of Change

1 Very much or much improved

2 Minimally improved

3 No change

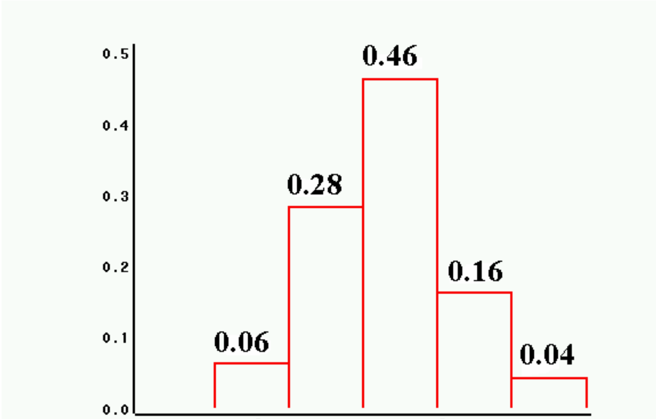
4 Minimally worse

5 Much or very much worse

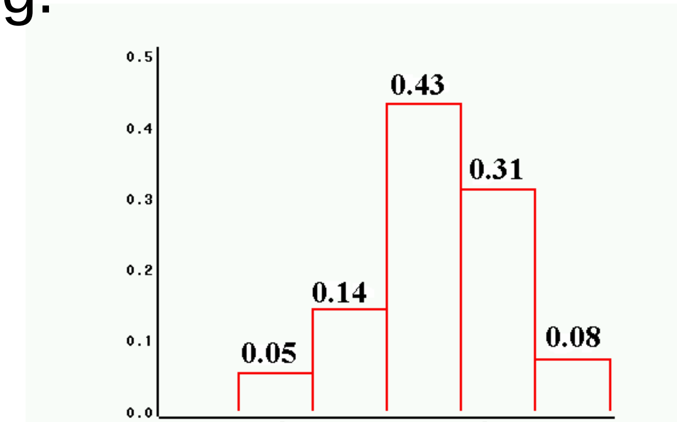
Stratum (Study)	Treatment	CGIC5 C1	C2	C3	C4	C5	Total
1	Tacrine	4	23	45	22	2	96
	placebo	2	22	54	29	3	110
2	Tacrine	14	119	180	54	6	373
	placebo	1	22	35	11	3	72
3	Tacrine	13	20	24	10	1	68
	placebo	7	16	17	10	3	53
4	Tacrine	21	106	175	62	17	381
	placebo	8	24	73	52	13	170
5	Tacrine	3	14	19	3	0	39
	placebo	2	13	18	7	1	41

Odds proportional between treatments within strata (studies), not between strata

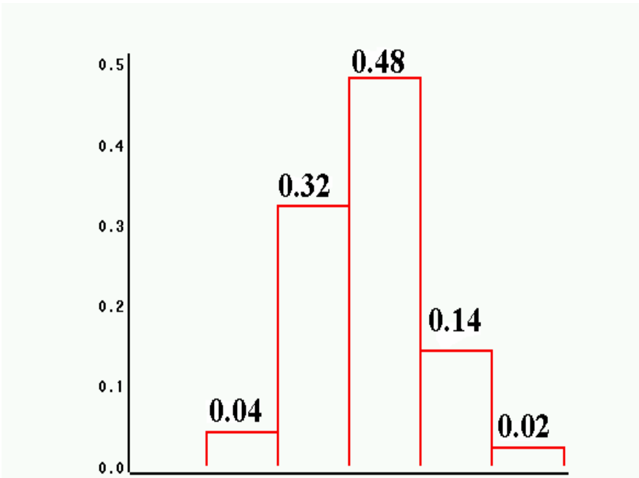
e.g.



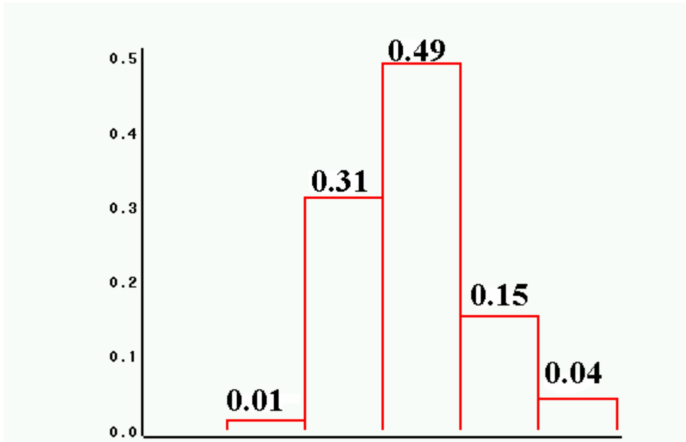
Study 4 Tacrine



Study 4 placebo



Study 2 Tacrine



Study 2 placebo

21 parameters [(k – 1) r + 1]:

Stratum 1:	α_{11},	α_{12},	α_{13},	α_{14}	20 strata split parameters
Stratum 2:	α_{21},	α_{22},	α_{23},	α_{24}	
Stratum 3:	α_{31},	α_{32},	α_{33},	α_{34}	
Stratum 4:	α_{41},	α_{42},	α_{43},	α_{44}	
Stratum 5:	α_{51},	α_{52},	α_{53},	α_{54}	

Treatment: θ

Fit model using SAS PROC NLMIXED *as shown in Supplement 6.4*

Note beta1 corresponds to θ

Estimation

$$\hat{\theta} = 0.505$$

$$se(\hat{\theta}) = 0.112$$

$$95\%\ \text{C.I. for } \hat{\theta}:$$

$$[\hat{\theta} \pm 1.96\, se\, (\hat{\theta})]$$

$$[0.505 \pm 1.96(0.112)]$$

$$[0.285,\ 0.725]$$

Hypothesis test

$$-2\text{Log}\hat{L} = 3559.7$$

To test $\theta = 0$ need to fit model without treat

Likelihood ratio test (H_0: $\theta = 0$)

$$= -2\text{Log}\hat{\text{L}}_{(\text{without treat})} - -2\text{Log}\hat{\text{L}}_{(\text{with treat})}$$

= 3580.1 – 3559.7 = 20.4

$$(\text{cf. } \chi_1^2)$$