# 17

# MODELLING ORDINAL AND MULTINOMIAL DATA

## OBJECTIVES

After reading this chapter, you should be able to:

1. Select an appropriate model from the following based upon the objectives of your study and the nature of your data
   - multinomial logistic model
   - proportional-odds model
   - adjacent-category model
   - continuation-ratio model.
2. Fit all of the models listed above.
3. Evaluate the assumptions on which the models are based and use one or more tests to compare different models.
4. Interpret *OR* estimates from each of the models.
5. Compute predicted probabilities from each of the models.

## 17.1 INTRODUCTION

In some studies, the outcome of interest might be categorical but have more than 2 categories (*ie* multinomial). These data could be recorded on either a nominal or ordinal scale. Nominal data arise when the outcome categories have no specific ordering (*eg* reason for culling might be classified as due to low production, reproduction, mastitis or other). Ordinal data arise when the outcome categories have a distinct order to them (*eg* severity of disease might be classified as absent, mild, moderate or severe). Clinical outcome data may better be analysed by treating the results as ordinal data rather than dichotomising the result (Norris *et al*, 2006; Valenta *et al*, 2006).

Nominal data can be analysed using log-linear models or multinomial logistic regression models. Log-linear models can simultaneously evaluate the effects of multiple predictors on multiple outcomes but are limited to the evaluation of categorical variables (predictors and outcomes). Log-linear models are used less frequently than regression-type models in veterinary epidemiology so they will not be discussed further.

An overview of a variety of regression models applicable to nominal and ordinal data is presented in Section 17.2. Each of the 4 models introduced in that section is described in more detail in Sections 17.3 to 17.7. All of the examples used in this chapter are based on data derived from a study designed to determine if ultrasound evaluation of beef cattle at the start of the feeding (fattening) period could be used to predict whether the carcass from the animal would eventually be graded as 1=AAA (highest grade), 2=AA, or 3=A (lowest grade in terms of price) (Keefe *et al*, 2004). This classification is based on the amount of 'marbling' (intramuscular fat in the loin region) present in the carcass after slaughter with grade AAA selling for the highest price. The dataset (beef_ultra) is described more fully in Chapter 31, but the main variables used in this chapter are shown in Table 17.1.

**Table 17.1 Variable from beef ultrasound dataset (beef_ultra)**

| | |
|---|---|
| farm | farm id |
| id | animal id |
| grade | carcass grade 1=AAA 2=AA 3=A<br>(AAA is the "best" grade) |
| bckgrnd | 'backgrounding' (animal spends time on pasture between weaning and entering the feedlot) (0=no, 1=yes) |
| sex | 0=heifer (female)<br>1=steer (castrated male) |
| backfat | backfat thickness (mm) |
| ribeye | area of ribeye muscle (sq cm) |
| imfat | intramuscular fat score (%) |
| carc_wt | carcass weight (kg) |

## 17.2 OVERVIEW OF MODELS

An overview of the 4 models to be discussed in this chapter is presented here. In each case we will assume that the outcome has $J$ categories with $j$ being used to designate the categories from 1 to $J$ (*ie* $j$=1,...,$J$). For the sake of simplicity, we will assume that there is a single dichotomous predictor in the model, but these models can easily be extended to multiple predictors. A simple example, based on the data in Table 17.2, will be used to demonstrate most of the models. All models discussed in this Chapter are presented as logistic models; they can be fit as other binomial models (*eg* probit, complementary log log) but these are beyond the scope of this text. More details about these models can be found in (Hilbe, 2009; Long, 1997; Long & Freese, 2006).

**Table 17.2 Cross-tabulation of grade and backgrounding from dataset beef_ultra**

| Category | Grade | Backgrounded | Not backgrounded | Totals |
|---|---|---|---|---|
| 1 | AAA | 149 | 15 | 164 |
| 2 | AA | 198 | 79 | 277 |
| 3 | A | 20 | 26 | 46 |
| | | 367 | 120 | 487 |

### 17.2.1 Multinomial logistic model

Nominal data can be analysed using a **multinomial logistic model** which relates the probability of being in category $j$ to the probability of being in a baseline category (which we will refer to as category 1). The model can be written as follows.

$$\ln\frac{p(Y=j)}{p(Y=1)}=\beta_0^{(j)}+\beta_1^{(j)}X \qquad \textit{Eq 17.1}$$

A complete set of coefficients ($\beta_0$ and $\beta_1$) is estimated for each of the $J$-1 levels being compared with the baseline (these are designated as $\beta^{(j)}$). Graphically, the effect of the predictor can be seen in Fig. 17.1.

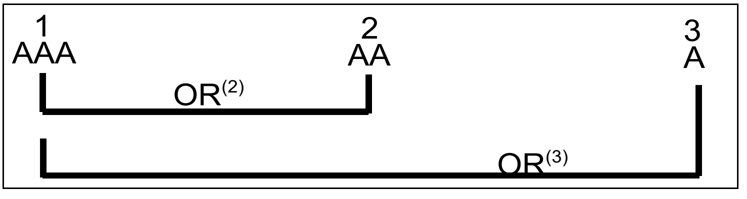



**Fig. 17.1 Multinomial logistic model**

Based on the data in Table 17.2, the odds ratio ($OR$) for a backgrounded animal being in category 2 (AA) (compared with category 1) is:

$$OR^{(2)}=\frac{15*198}{149*79}=0.25$$

Similarly, the $OR$ for category 3 (A) compared with category 1 (AAA) is:

$$OR^{(3)} = \frac{15*20}{26*149} = 0.08$$

### 17.2.2 Proportional odds model

The multinomial model does not make any assumptions about the ordering of the categories. An approach for analysing ordinal data is to use a proportional odds model which relates the probability of being at or above a category to the probability of being in any lower category. The model assumes that this relationship is the same at each of the categories. The model can be written as follows.

$$\ln \frac{p(Y \geq j)}{p(Y < j)} = \beta_0^{(j)} + \beta_1 X$$ ***Eq 17.2***

Fitting this model requires that $J$-1 intercepts ($\beta_0$) be estimated, but only a single $\beta_1$. Graphically, the effects of the predictor can be seen in Fig. 17.2.

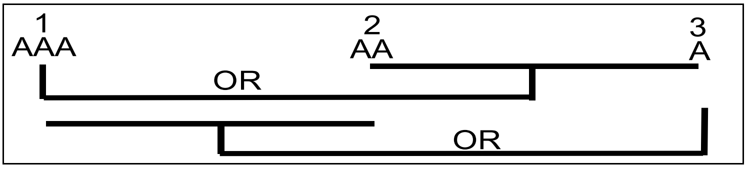



**Fig. 17.2 Proportional odds model**

Based on the data in Table 17.2, the *OR* associated with being backgrounded for category 2 or 3 (compared with category 1) is:

$$OR^{(2)} = \frac{(198+20)*15}{(79+26)*149} = 0.21$$

while the *OR* associated with being backgrounded for category 3 (compared with being <3) (*ie* A vs AA or AAA) is:

$$OR^{(3)} = \frac{(15+79)*20}{(149+198)*26} = 0.21$$

Because the 2 *OR*s are the same, the assumption of proportional odds seems to hold. However, this may not be true for all of the predictors we are interested in.

### 17.2.3 Adjacent-category model

If the categories are ordered, and in some sense 'equidistant', then a constrained multinomial model, or **adjacent-category model** can be fit to the data. This model is based on the assumption that the predictor increases (or decreases) the log odds of a category occurring by a fixed amount as you go up through the categories. Consequently, the model can be written as follows.

$$\ln \frac{p(Y = j)}{p(Y = j-1)} = \beta_0^{(j)} + (J-1)\beta_1 X$$ ***Eq 17.3***

Fitting this model requires that *J*-1 intercepts ($\beta_0$) be estimated, but only a single $\beta_1$. Graphically, the effects of the predictor can be seen in Fig. 17.3.

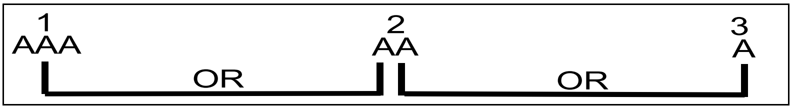



**Fig. 17.3 Adjacent-category model**

The estimate of $\beta_1$ cannot be derived easily from the data in Table 17.2, but the *OR* for AA vs AAA is 0.276 while that for A vs AAA is $(0.276)^2$=0.076.

### 17.2.4 Continuation-ratio model

An alternative for analysing ordinal data is to use a continuation-ratio model which relates the probability of being in a category to the probability of being in any lower category. The model can be written as follows.

$$\ln \frac{p(Y=j)}{p(Y<j)} = \beta_0^{(j)} + \beta_1^{(j)} X \qquad \textbf{\textit{Eq 17.4}}$$

A complete set of coefficients ($\beta_0$ and $\beta_1$) is estimated for each of the *J*-1 categories above the baseline. Graphically, the effect of the predictor can be seen in Fig. 17.4.

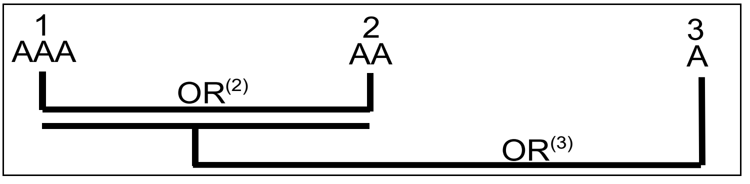



**Fig. 17.4 Continuation-ratio model**

Based on the data in Table 17.2, the *OR* associated with being backgrounded for category 2 (compared with category 1) is:

$$OR^{(2)} = \frac{15*198}{149*79} = 0.25$$

while the *OR* associated with being backgrounded for category 3 (compared with being <3) (*ie* A vs AA or AAA) is:

$$OR^{(3)} = \frac{(15+79)*20}{(149+198)*26} = 0.21$$

## 17.3 MULTINOMIAL LOGISTIC REGRESSION

In multinomial logistic regression of an outcome that has *J* categories, the probability of membership in each of the outcome categories is computed by simultaneously fitting *J*-1 separate logistic models (with one category serving as the baseline or reference category). Consequently, for a dependent variable with 4 levels (leaving the first level as the baseline category), we estimate 3 sets of coefficients ($\beta^{(2)}$, $\beta^{(3)}$, $\beta^{(4)}$) corresponding to the remaining outcome categories. Because $\beta^{(1)}=0$, the predicted probability that an observation is in category 1 is:

$$p(Y=1)=\frac{1}{1+\exp(X\beta^{(2)})+\exp(X\beta^{(3)})+\exp(X\beta^{(4)})} \qquad \textbf{\textit{Eq 17.5}}$$

while the probability of being in category 2 is:

$$p(Y=2)=\frac{\exp(X\beta^{(2)})}{1+\exp(X\beta^{(2)})+\exp(X\beta^{(3)})+\exp(X\beta^{(4)})} \qquad \textbf{\textit{Eq 17.6}}$$

and similarly for categories 3 and 4.

### 17.3.1 Odds ratios

For any given predictor (*eg* -bckgrnd-), there is a separate estimate of the effect of that predictor on each outcome (relative to the base level). Exponentiation of the coefficients from a multinomial regression model produces odds ratios as a measure of effect. **Note** Strictly speaking, these effect measures are not odds ratios. They are actually the ratio of 2 relative risks (or risk ratios) with each relative risk describing the probability of the outcome in the category of interest relative to the baseline category. Consequently, it would be more appropriate to refer to them as relative risk ratios and some computer programs do so. However, the term odds ratio is commonly applied and will be used in this chapter.

Example 17.1 shows a very simple model for carcass classification with -bckgrnd- as the single predictor. The odds ratios are exactly the same as were found in Section 17.2. They indicate that an animal that was backgrounded was 0.25 times as likely to grade AA (compared with AAA) as an animal that was not backgrounded. Similarly, backgrounded animals were 0.08 times as likely as a non-backgrounded animal to grade A.

Both of the *OR*s in Example 17.1 suggest that backgrounding decreased the risk of a lower carcass grade and this effect was clearly statistically significant (see Section 17.3.3 for assessment of significance).

As with ordinary logistic regression, multinomial logistic regression can be extended to model the effects of multiple predictors that might be categorical or continuous in nature. Example 17.2 shows a model for carcass grade including predictors with results presented as coefficients.

### 17.3.2 Interpretation of coefficients

Estimates (coefficients or *OR*s) from multinomial logistic regression models are interpreted in a manner similar to those from ordinary logistic regression. The *OR* for the predictor -imfat- in Example 17.2 suggests that, for a unit increase in the intramuscular fat reading from the ultrasound examination, the odds of being downgraded from AAA to AA goes down by a factor of $e^{-0.444}=0.64$ (36% reduction) while the odds of being graded A goes down by a factor of $e^{-1.041}=0.35$ (65% reduction). In Example 17.2, all of the predictors have more pronounced

**Example 17.1 Simple multinomial logistic regression**
data = beef_ultra

A simple multinomial logistic regression of carcass grade (3 levels) was carried out with -bckgrnd- as the sole predictor. Carcass grade AAA was the baseline (referent) level.

The first table presents the results in terms of coefficients of the logistic models.

Number of obs = 487
LR chi2 (2) = 49.33
Prob > chi2 = <0.001
Log likelihood = -418.670

| | | Coef | SE | Z | P | 95% CI | |
|---|---|---|---|---|---|---|---|
| **AA** | | | | | | | |
| | **bckgrnd** | -1.377 | 0.302 | -4.560 | 0.000 | -1.969 | -0.786 |
| | **constant** | 1.661 | 0.282 | 5.900 | 0.000 | 1.109 | 2.213 |
| **A** | | | | | | | |
| | **bckgrnd** | -2.558 | 0.402 | -6.360 | 0.000 | -3.347 | -1.770 |
| | **constant** | 0.550 | 0.324 | 1.700 | 0.090 | -0.085 | 1.186 |

Being a backgrounded reduced the logit of the probability of grading AA or A by 1.38 and 2.56 units, respectively.

The second table presents the results in terms of odds ratios.

| | | OR | SE | 95% CI | |
|---|---|---|---|---|---|
| **AA** | | | | | |
| | **bckgrnd** | 0.252 | 0.076 | 0.140 | 0.456 |
| **A** | | | | | |
| | **bckgrnd** | 0.077 | 0.031 | 0.035 | 0.170 |

Backgrounded animals were 0.25 and 0.08 times as likely to be downgraded to AA or A compared with non-backgrounded animals.

effects on the A vs AAA comparison compared with the AA vs AAA comparison. This was expected given the ordinal nature of the data, but nothing in the model guarantees this. This pattern would not be expected if unordered nominal data were being analysed.

### 17.3.3 Testing significance of predictors

The significance of predictors can be assessed using either a Wald test or a likelihood ratio test (*LRT*). Overall tests of significance can be carried out (for all logistic models fit) as well as tests for coefficients within individual models. Note however, that tests of significance for a predictor within a given logistic model (*eg* for grade=A) will change if the baseline category is changed. Consequently, overall tests of significance provide a better estimate of the significance of the predictor. Unconditionally, -sex- was not a significant predictor (Wald test P-value = 0.46—data not shown). However, based on the model in Example 17.2, while the

**Example 17.2 Multiple multinomial logistic regression**
data = beef_ultra

Prediction of carcass grade based on the background status, sex and weight of the animal and three ultrasound measurements determined at the start of the feeding period.

Number of obs = 487
LR chi2 (10) = 146.08
Prob > chi2 < 0.001
Log likelihood = -370.298

| | | Coef | SE | Z | P | 95% CI | |
|---|---|---|---|---|---|---|---|
| **AA** | | | | | | | |
| | **bckgrnd** | -1.282 | 0.337 | -3.800 | 0.000 | -1.944 | -0.621 |
| | **sex** | 0.906 | 0.266 | 3.400 | 0.001 | 0.384 | 1.429 |
| | **backfat** | -0.249 | 0.116 | -2.140 | 0.032 | -0.477 | -0.021 |
| | **ribeye** | 0.373 | 0.081 | 4.620 | 0.000 | 0.215 | 0.532 |
| | **imfat** | -0.444 | 0.123 | -3.610 | 0.000 | -0.684 | -0.203 |
| | **carc_wt** | -0.019 | 0.004 | -5.490 | 0.000 | -0.026 | -0.012 |
| | **constant** | 6.806 | 1.276 | 5.330 | 0.000 | 4.304 | 9.307 |
| **A** | | | | | | | |
| | **bckgrnd** | -1.830 | 0.483 | -3.790 | 0.000 | -2.777 | -0.884 |
| | **sex** | 1.486 | 0.458 | 3.250 | 0.001 | 0.589 | 2.384 |
| | **backfat** | -0.667 | 0.258 | -2.580 | 0.010 | -1.173 | -0.161 |
| | **ribeye** | 0.493 | 0.152 | 3.240 | 0.001 | 0.195 | 0.791 |
| | **imfat** | -1.041 | 0.236 | -4.400 | 0.000 | -1.504 | -0.578 |
| | **carc_wt** | -0.040 | 0.007 | -6.160 | 0.000 | -0.053 | -0.027 |
| | **constant** | 13.976 | 2.221 | 6.290 | 0.000 | 9.624 | 18.329 |

Carcass grade=AAA was the baseline (referent) level.

Once again, -bckgrnd- significantly reduces the risk of lower grades (AA and A). -carc_wt-, -backfat-, -ribeye- and -imfat- are intervening variables between -bckgrnd- and -grade- so the estimate of effect of -bckgrnd- is just the direct effect. See Chapter 13 for a discussion of intervening variables.

Wald and likelihood ratio tests for -sex- give slightly different values ($\chi^2$ of 15.0 and 15.5, respectively on 2 df); both were highly significant (P<0.001). Control of other factors (intervening variables), has made sex an important predictor of carcass grade.

### 17.3.4 Obtaining predicted probabilities

Predicted probabilities of the occurrence of each outcome category can be computed from the multinomial logistic regression (see Eqs 17.5 and 17.6). These will, of course, vary with the values of the predictors for the animal. Table 17.3 shows those values for a small subset of the cattle based on the model shown in Example 17.2.

**Table 17.3 Predicted probabilities from a multinomial logistic regression model**

| | | | | | | | | probability of grade | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **id** | **grade** | **bckgrnd** | **sex** | **backfat** | **ribeye** | **imfat** | **carc_wt** | **AAA** | **AA** | **A** |
| 1 | AA | bckg | steer | 2.5 | 8.9 | 4.5 | 357.7 | 0.03 | 0.55 | 0.42 |
| 2 | AA | bckg | steer | 5.9 | 11.8 | 5.2 | 323.2 | 0.02 | 0.68 | 0.30 |
| 3 | AAA | bckg | steer | 3.1 | 9.7 | 3.5 | 360.0 | 0.05 | 0.65 | 0.30 |
| 4 | AA | bckg | female | 2.5 | 7.5 | 5.2 | 307.3 | 0.02 | 0.37 | 0.62 |
| 5 | AAA | bckg | steer | 1.9 | 8.0 | 4.9 | 354.5 | 0.03 | 0.48 | 0.49 |

The sum of the probabilities for each animal equals 1.

### 17.3.5 Assumption of independence of irrelevant alternatives (IIA)

The multinomial regression model is based on an assumption that the odds of one level of the outcome being observed is independent of what other alternatives are available. For the carcass data discussed, this would mean that if the odds of an AA were twice those of an A, they should always be twice, regardless of whether there were no alternatives or if the alternatives consisted of just AAA or also included other levels (B, C *etc*).

Two of the most commonly used tests of this assumption are the Hausman & McFadden (1984) and Small-Hsiao (1985) tests of IAA. Both are based on the principle of fitting a full model and comparing the coefficients from that model to a model with one or more of the alternatives deleted (partial model). The null hypothesis is that the coefficients from the full model are the same as from the partial model. If the P-value of the the test is >0.05 there is insufficient evidence to reject the null hypothesis (*ie* the assumption has been met). For the Hausman test, the statistic may be negative which is also assumed to support the null hypothesis.

Unfortunately, the two tests often give conflicting results and recent simulation studies (cited in Long & Freese (2006)) suggest that they may be of limited use in determining whether the assumption has been met. In the face of conflicting results, the best advice may be from the early statement of McFadden cited in Long & Freese (2006)) that multinomial models should only be used when the alternatives "can plausibly be assumed to be distinct and weighted independently in the eyes of the decision maker". For the beef carcass data, it seems unlikely that the effects of factors on the choice of a grade by an inspector would be independent of the range of choices available. Example 17.3 contains the results of these two tests for the beef carcass data.

**Example 17.3 Evaluating assumption of independence of irrelevant alternatives (IIA)**
data = beef_carcass

The P-values for the Hausman test of IIA were 0.768 and 0.993 if levels AA or A were left out respectively. Both values strongly support the notion that the assumption of IIA had been satisfied. The Small-Hsiao test of IIA produces different estimates each time it is run (due to a random element in the calculations) and the results were very unstable.

A likelihood ratio test of whether or not any of the levels could be combined produced P-values <0.001 for all pairwise combinations of levels, suggesting that no pairs of outcome levels could be combined.

It is also possible to statistically evaluate (using a Wald or likelihood ratio test) whether any of the outcome levels are not significantly different from other levels. If some are not, one might want to consider combining those levels. (See Example 17.3.)

### 17.3.6 Regression diagnostics

Specialised diagnostics for multinomial logistic regression are not as readily available as they are for ordinary logistic regression. One approach is to fit ordinary logistic models for pairs of comparisons (*eg* grade=A vs AAA and AA vs AAA) and evaluate the regression diagnostics for those models. An overall goodness-of-fit test has recently been developed (Fagerland *et al*, 2008) but at the time of writing was not readily available in standard software packages.

### 17.3.7 Models for outcomes with alternative specific data

In the beef carcass data example, none of the predictors vary across outcome alternatives (*ie* the carcass weight of the animal was constant, regardless of whether outcome A, AA or AAA was being considered). This is not always the case. Consider the situation in which a dog owner has to choose among 3 options for dealing with a recently diagnosed case of cancer in their dog. The options might include: treatment at their local clinic, treatment at a private referral hospital or treatment at a university based teaching hospital. Factors which might influence their decision might include: the age of the dog, their income level and the distance to the various clinics. While the first two factors (age and income) are independent of the alternatives, the last (distance) varies with the alternatives being considered (*eg* the local clinic is closer than the other alternatives). Various approaches for dealing with this situation exist (one of which is conditional logistic regression—Section 16.15) and the reader is referred to (Hilbe, 2009; Long & Freese, 2006) for an explanation of how to structure the data and fit an appropriate model for this situation.

## 17.4 MODELLING ORDINAL DATA

Ordinal data can arise in a variety of ways. For example, an observed continuous variable might be divided into categories. Alternatively, levels of an ordinal variable could represent categories of an unobserved (but hypothesised) continuous variable (*eg* opinions ranging from strongly agree to strongly disagree, or disease severity ranging from absent to severe). Finally, categories might represent total values of a composite variable made up of a series of scored variables (*eg* a hygiene score that represents the sum of scores from several questions about hygiene in a barn).

While the multinomial models described above can also be used to analyse ordinal data, they ignore the fact that the categories fall in a logical, ordered sequence. There are a number of different ways of fitting ordinal models. We will consider 3 of them: proportional-odds models (Section 17.5), adjacent-category models (Section 17.6) and continuation-ratio models (Section 17.7).

## 17.5 PROPORTIONAL ODDS MODEL (CONSTRAINED CUMULATIVE LOGIT MODEL)

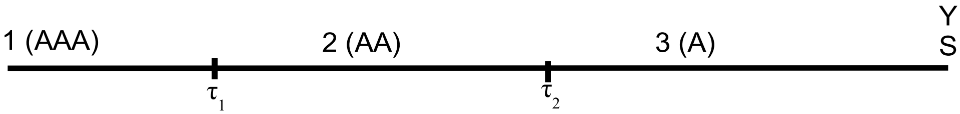
This is the most commonly encountered type of ordinal logistic model. In a proportional odds model, the coefficients measure the effect of a predictor on the log odds of being at or above a specified level compared with the log odds of being below the specified level. It is based on the assumption that the coefficients do not depend upon the outcome level, so only a single coefficient for each predictor is estimated. A graphic representation of this model is presented in Fig. 17.2.

A proportional odds model assumes that the ordinal outcome variable represents categories of an underlying continuous latent (unobserved) variable. Assume that the value of the underlying latent variable (or 'score') ($S_i$) is a linear combination of predictor variables.

$$S_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \varepsilon_i \qquad \textbf{\textit{Eq 17.7}}$$

where $\varepsilon_i$ is a random error term from a continuous distribution.

The latent variable ($S$) is divided by cutpoints ($\tau_j$) so that the $i^{th}$ individual is classified as category 1 (AAA) if $S_i \leq \tau_1$ and is classified as category 2 (AA) if $\tau_1 < S_i \leq \tau_2$, and so on.




The probability of observing outcome $j$ in the $i^{th}$ individual is:

$$p(\text{outcome}_i = j) = p(\tau_{j-1} < S_i < \tau_j) \qquad \textbf{\textit{Eq 17.8}}$$

If the random term ($\varepsilon_i$) is assumed to have a logistic distribution (with a mean of 0 and a variance of $\pi^2/3$), then

$$\text{p}(S_i < \tau_j) = \frac{1}{1 + \text{e}^{S_i - \tau_j}} \qquad \textbf{\textit{Eq 17.9}}$$

**Note** Assuming the latent variable has a normal distribution gives rise to an ordinal probit model, but those are not discussed in this chapter.

The model fit by assuming a logistically distributed latent variable can also be written as (presented with a single predictor $X$ for simplicity):

$$\text{logit}(\text{p}(Y \leq j)) = \beta_{0j} + \beta X$$

where the $\beta_{0j}$ are intercepts and $\beta$ is the effect (slope) of the predictor. Thus, the model is one in which the log odds of the outcome can be viewed as being represented by a series of parallel lines with different intercepts.

Example 17.4 presents a proportional-odds model for the carcass grade data.

**Example 17.4 Proportional-odds model**
data = beef_ultra

A proportional-odds model was fit to the carcass grade data with the same predictors as used in Example 17.2 and 17.3.

Number of obs = 487
LR chi2 (6) = 138.53
Prob > chi2 < 0.001
Log likelihood = -374.070

| | Coef | SE | Z | P | 95% CI | |
|---|---|---|---|---|---|---|
| **bckgrnd** | -1.214 | 0.263 | -4.620 | 0.000 | -1.729 | -0.699 |
| **sex=steer** | 0.862 | 0.231 | 3.730 | 0.000 | 0.410 | 1.315 |
| **backfat** | -0.287 | 0.106 | -2.690 | 0.007 | -0.495 | -0.078 |
| **ribeye** | 0.335 | 0.070 | 4.790 | 0.000 | 0.198 | 0.472 |
| **imfat** | -0.520 | 0.109 | -4.760 | 0.000 | -0.734 | -0.306 |
| **carc_wt** | -0.022 | 0.003 | -7.140 | 0.000 | -0.028 | -0.016 |
| **cutpoint 1** | -8.635 | 1.105 | | | -10.801 | -6.470 |
| **cutpoint 2** | -4.928 | 1.038 | | | -6.962 | -2.894 |

The odds ratio associated with being a steer, compared with being a female is:

$$e^{0.862}=2.37$$

This suggests that being a steer increases the odds of being at or above any given carcass grade compared with being below that grade by 2.37 times. (Remember that the data are coded so that A is grade 3—*ie* greatest economic loss). As such it measures the overall increased chance of a poor (higher) grade that is associated with being a steer.

### 17.5.1 Predicted probabilities

The first observation in the dataset is a backgrounded steer (sex=1) with a backfat=2.51, a ribeye=8.94, an imfat=4.46 and a carc_wt=357.7. For this animal, the latent variable ($S_1$) is:

$$S_i=-8.329$$

Consequently, the probability of this animal being in category 1 (AAA) (from Eq 17.9) is:

$$p(Y=1)=\frac{1}{1+e^{-8.329-(-8.635)}}=0.42$$

Similarly, the probability of this animal being graded AA is 0.54 and A is 0.03.

The probabilities of each grade outcome for the first 5 animals from this dataset (and the values of the predictor variables for those animals) are shown in Table 17.4.

**Table 17.4 Values of predictor variables, latent variables ($S_i$) and predicted probabilities of each of the carcass grades from the proportional-odds model**

| | | | | | | | | | probability of grade | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **id** | **grade** | **bckgrnd** | **sex** | **backfat** | **ribeye** | **imfat** | **carc_wt** | **S** | **AAA** | **AA** | **A** |
| 1 | AA | bckg | steer | 2.5 | 8.9 | 4.5 | 357.7 | -8.33 | 0.42 | 0.54 | 0.03 |
| 2 | AA | bckg | steer | 5.9 | 11.8 | 5.2 | 323.2 | -7.99 | 0.34 | 0.61 | 0.05 |
| 3 | AAA | bckg | steer | 3.1 | 9.7 | 3.5 | 360.0 | -7.82 | 0.31 | 0.64 | 0.05 |
| 4 | AA | bckg | fem. | 2.5 | 7.5 | 5.2 | 307.3 | -8.93 | 0.57 | 0.41 | 0.02 |
| 5 | AAA | bckg | steer | 1.9 | 8.0 | 4.9 | 354.5 | -8.60 | 0.49 | 0.48 | 0.03 |

The effect of a single predictor (-imfat-) on the predicted probability can best be viewed by generating smoothed curves of the probability of each grade against -imfat-. Fig. 17.5 shows a graph of lowess smoothed probabilities (smoothed with a bandwidth of 50%) of each grade against the intramuscular fat level (over the range of 3.0 to 6.0—the range in which most -imfat- values fall). **Note** As the probability of each outcome depends on the value of all predictors in the model, the smoothed curves shown in Fig. 17.5 represent average probabilities of the grade as -imfat- changes.

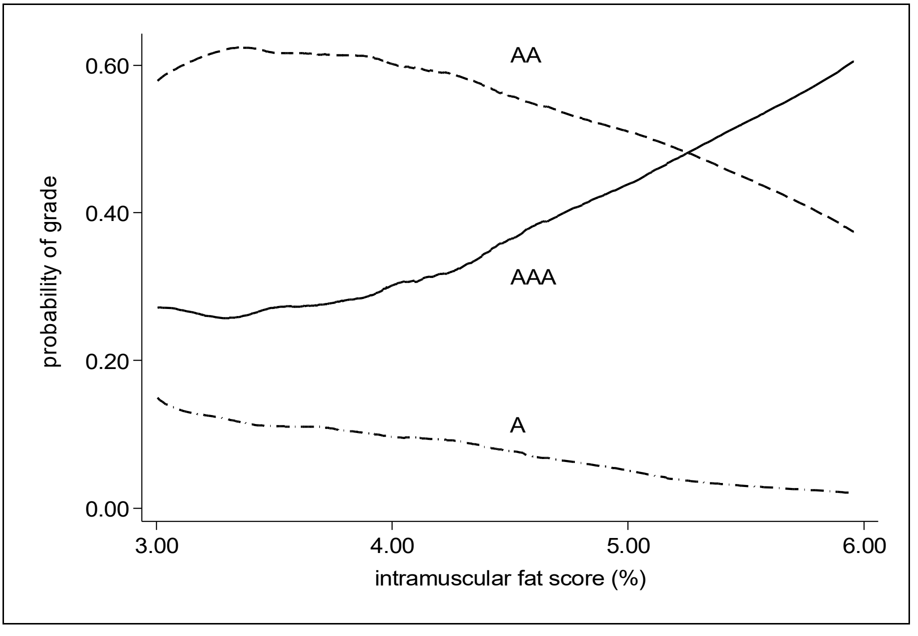



**Fig. 17.5 Smoothed mean probabilities of grades**

As can be seen, the probability of a carcass being graded AA or A goes down as the intramuscular fat level at the start of the feeding period goes up. On the other hand, the probability of a grade of AAA goes up substantially.

### 17.5.2 Evaluating the proportional-odds assumption

A rough assessment of the assumption of proportional odds can be obtained by comparing the log likelihood of the ordered logit model ($L_1$) with one obtained from the multinomial logit model ($L_0$) using a likelihood ratio test. If there are $k$ predictors (not counting the intercept) and $J$ categories of outcome, the multinomial model will fit $(k+1)(J-1)$ parameters, while the proportional-odds model will fit $k+(J-1)$ so the difference in degrees of freedom is $k(J-2)$. Consequently, $-2(L_1-L_0)$ should have an approximate $\chi^2$ distribution with $k(J-2)$ degrees of freedom. **Note** This is only an approximate test because the proportional-odds model is not nested within the multinomial model. However, it gives a rough assessment of the assumption of the proportional-odds assumption.

In our example, the log likelihoods of the multinomial and proportional odds models were -370.30 and -374.07, respectively so the *LRT* is:

$$LRT=-2(-370.30-[-374.07])=7.54$$

The $\chi^2$ statistic has $k(J-2)=6$ df which yields a P-value of 0.27. Consequently, there is no evidence that the proportional-odds assumption does not hold. As an alternative to comparing the ordinal logistic model to a multinomial model, the comparison can be made with a generalised ordinal logistic model (described below—Section 17.5.3). This comparison yields a $\chi^2$ of 6.50 (P=0.37).

An alternative approximate *LRT* based on fitting $J$-1 separate binary models has been developed (Wolfe & Gould, 1998). The models are fit first assuming the $\beta$s are constant across all models (proportional-odds assumption) and the sum of these log likelihoods are compared with the sum of those obtained by fitting the models without the assumption of constant $\beta$s. For the beef ultrasound model, this test produces a $\chi^2$ value of 6.42 (P=0.38).

The likelihood ratio tests described above are omnibus tests which evaluate the assumption of proportional odds over all predictors. A Wald test which will provide an overall assessment as well as an evaluation of the assumption for each predictor separately is available (Brant, 1990). The results of this test for the model fit in Example 17.4 are presented in Table 17.5.

**Table 17.5 Brant (Wald) test of proportional-odds assumption**

| Variable | $X^2$ | P | df |
|---|---|---|---|
| all | 6.66 | 0.35 | 5 |
| bckgrnd | 1.42 | 0.23 | 1 |
| sex | 0.18 | 0.67 | 1 |
| backfat | 0.59 | 0.44 | 1 |
| ribeye | 1.47 | 0.23 | 1 |
| imfat | 0.80 | 0.37 | 1 |
| carc_wt | 0.43 | 0.51 | 1 |

The P-value of the overall Wald test is comparable to the last two approximate likelihood ratio tests described above. None of the individual predictors have significant test results suggesting that the proportional-odds assumption is valid. Other tests of the proportional-odds assumption are available but there are no clear guidelines for choosing one test over another. In general, if any of the tests discussed above yields a significant result, the assumption should be investigated further.

### 17.5.3 Dealing with non-proportional odds

In the event that one or more predictors appears to violate the assumption of proportional odds, there are a number of potential approaches to dealing with the problem. A **generalised ordinal logistic regression model** is one in which a complete set of coefficients are estimated for each cutpoint in the ordinal model (*eg* A vs AA/AAA and A/AA vs AAA). Consequently, it is no more parsimonious than the multinomial model, but it does take into account the ordering of the categories. The log-likelihood of this model can be compared with that of a model assuming proportional odds to see if the assumption is valid (see Section 17.5.2).

If the proportional odds assumption appears to hold for some predictors, but not all, it is possible to fit a **partial proportional odds model** in which the assumption of proportional odds is removed for selected predictors. For our example, there were no predictors which showed significant evidence of violating the proportional odds assumption (Table 17.5), but the 2 with the smallest P-values were -bckgrnd- and -ribeye-. If the coefficients for these two predictors are allowed to vary across cut-points, but the remainder are constrained to be constant (proportional odds), the log-likelihood for the model is -371.70 which yields a likelihood ratio test $\chi^2$ of 4.73 (P=0.09), providing some limited evidence that the effects of these 2 predictors might differ across the 2 cutpoints.

Two other approaches for dealing with non-proportional odds are the **stereotype logistic model** and the **heterogeneous choice logistic model**. These are beyond the scope of this text and the reader is referred to (Hilbe, 2009; Long & Freese, 2006) for details.

### 17.5.4 Regression diagnostics

As with multinomial models, regression diagnostics for ordinal models are not well developed. Hosmer and Lemeshow (2000) suggest fitting ordinary logistic models to data based on the cutpoints in the ordinal data (*eg* one model which compares A to AA/AAA and one which compares A/AA to AAA). Residuals from these models can be evaluated using techniques described in Chapter 16.

## 17.6 ADJACENT-CATEGORY MODEL

In an adjacent-category logistic regression model, each coefficient measures the effect of a factor on the logit of the probability of being in a specified level compared with the probability of being in the level below. For any given predictor, this results in the estimation of a single effect that expresses how the predictor influences the log odds of the outcome moving up to the next (adjacent) category. This model is also known as a constrained multinomial model because it is estimated as a multinomial model with the constraint that the coefficient for categories *n* levels apart be *n* times the coefficient for adjacent categories. (Alternatively, the *OR* for categories *n* levels apart will be the *OR* for adjacent levels raised to the power *n*.) This model is based on the assumption that, as you go from one level to the next, the *OR* is constant. A graphic representation is shown in Fig. 17.3.

Example 17.5 presents an adjacent-category model based on the multinomial model fit in Example 17.2. A likelihood ratio test can be used to compare this ‘constrained multinomial model’ with the usual multinomial model. If the test is significant, it suggests that the multinomial model is superior. The *LRT* for the model in Example 17.5 had a $\chi^2$ of 7.14 with 6 df (because 6 fewer coefficients were estimated) with a P-value of 0.31, suggesting that there is little evidence that the unconstrained model fits the data better than the adjacent-category model. In this case, for the sake of simplicity, the adjacent-category model is preferable.

**Example 17.5 Adjacent-category model**
data = beef_ultra

An adjacent-category model was fit using the same predictors presented in Example 17.2. The constraint that coefficients for categories two levels apart be twice those of the adjacent categories reduces the number of parameters which need to be estimated.

Number of obs = 487
LR chi2 (6) = 138.94
Prob > chi2 < 0.001
Log likelihood = -373.87

| | | Coef | SE | Z | P | 95% CI | |
|---|---|---|---|---|---|---|---|
| **AA** | | | | | | | |
| | **bckgrnd** | -1.040 | 0.235 | -4.420 | 0.000 | -1.501 | -0.579 |
| | **sex=steer** | 0.784 | 0.210 | 3.730 | 0.000 | 0.372 | 1.196 |
| | **backfat** | -0.276 | 0.097 | -2.850 | 0.004 | -0.465 | -0.086 |
| | **ribeye** | 0.296 | 0.063 | 4.670 | 0.000 | 0.172 | 0.420 |
| | **imfat** | -0.477 | 0.101 | -4.740 | 0.000 | -0.675 | -0.280 |
| | **carc_wt** | -0.020 | 0.003 | -6.970 | 0.000 | -0.026 | -0.014 |
| | **constant** | 7.844 | 1.036 | 7.570 | 0.000 | 5.813 | 9.876 |
| **A** | | | | | | | |
| | **bckgrnd** | -2.080 | 0.471 | -4.420 | 0.000 | -3.002 | -1.157 |
| | **sex=steer** | 1.568 | 0.420 | 3.730 | 0.000 | 0.744 | 2.391 |
| | **backfat** | -0.552 | 0.194 | -2.850 | 0.004 | -0.931 | -0.172 |
| | **ribeye** | 0.592 | 0.127 | 4.670 | 0.000 | 0.343 | 0.840 |
| | **imfat** | -0.955 | 0.202 | -4.740 | 0.000 | -1.350 | -0.560 |
| | **carc_wt** | -0.040 | 0.006 | -6.970 | 0.000 | -0.051 | -0.029 |
| | **constant** | 12.462 | 1.967 | 6.340 | 0.000 | 8.607 | 16.317 |

Outcome grade=AAA is the comparison group.

**Note** The coefficient for each predictor for grade=A is twice that for grade=AA because it is two categories away from AAA. For example, being a steer increases the log odds of being graded A by 1.57 units but the log odds of being graded AA by only 0.78 units.

## 17.7 CONTINUATION-RATIO MODEL

In continuation-ratio models, the log *OR* measures the effect of a factor on the odds of being in a specified level compared with the odds of being in any of the lower levels. This type of model is useful in situations where the dependent variable represents the number of attempts required to achieve an outcome (*eg* number of breedings required for conception in dairy cows). The individual must pass through all lower levels to reach the current level (you can't have your $3^{rd}$ breeding until you have had your $1^{st}$ and $2^{nd}$), hence the name 'continuation-ratio'. A graphic representation is shown in Fig. 17.4.

This model can be fit as a series of simple logistic models in which the dependent variable ($Y$) is recoded to be 1 for the level of interest, 0 for all lower levels and missing for all higher levels. For example, a continuation-ratio model evaluating the effects of predictors on the probability of conception for up to 4 breedings would require 3 separate logistic regressions. The data would be recoded as shown in Table 17.6.

**Table 17.6 Coding of data for a continuation-ratio model of effect of predictors on number of services required for conception**

| | Breeding | | | |
|---|---|---|---|---|
| | **1** | **2** | **3** | **4** |
| **Y1** | 0 | 1 | missing | missing |
| **Y2** | 0 | 0 | 1 | missing |
| **Y3** | 0 | 0 | 0 | 1 |

In this example, the coefficient for a predictor represents the effect of the factor on the log odds of conceiving on the $j^{th}$ breeding, conditional on not conceiving on any previous breedings.

The model contains the same number of parameters as the multinomial model presented in Section 17.3. Consequently, the model is no more 'parsimonious', but it results in estimates of the *OR* which have different interpretations than those from a multinomial logistic regression model. A constrained continuation-ratio model can be fit with the *OR* for each predictor constrained to be equal for each increment in the outcome. A likelihood ratio test, comparing the constrained and unconstrained models, can be used to evaluate the assumption of equal *OR*s.

The *OR* from the separate logistic models for the beef ultrasound data are not presented because it does not make biological sense to fit these data with a continuation-ratio model (*ie* movements between grades are not sequential events).

## References

Brant R. Assessing proportionality in the proportional odds model for ordinal logistic regression Biometrics. 1990; 46: 1171-8.

Fagerland MW, Hosmer DW, Bofin AM. Multinomial goodness-of-fit tests for logistic regression models Stat Med. 2008; 27: 4238-53.

Hausman J, McFadden D. Specification tests for the multinomial logit model Econometrica. 1984; 52: 1219-40.

Hilbe J. Logistic Regression Models. CRC Press; Boca Raton FL. 2009.

Hosmer D, Lemeshow S. Applied Logistic Regression. 2nd Ed. John Wiley and Sons; New York. 2000.

Keefe G, Dohoo I, Valcour J, Milton R. Ultrasonic imaging of marbling at feedlot entry as a predictor of carcass quality grade Canadian J Anim Sci. 2004; 84: 165-70.

Long J. Regression Models for Categorical and Limited Dependent Variables. Sage Publications; Thousand Oaks CA. 1997.

Long J, Freese J. Regression Models for Categorical Dependent Variables Using Stata. 2nd Ed. Stata Press; College Station, TX. 2006.

Norris CM, Ghali WA, Saunders LD, Brant R, Galbraith D, Faris P, Knudtson ML. Ordinal regression model and the linear regression model were superior to the logistic regression models J Clin Epidemiol. 2006; 59: 448-56.

Small K, Hsiao C. Multinomial logit specification tests International Economic Review. 1985; 26: 619-27.

Valenta Z, Pitha J, Poledne R. Proportional odds logistic regression - effective means of dealing with limited uncertainty in dichotomizing clinical outcomes Stat Med. 2006; 25: 4227-34.

Wolfe R, Gould W. An approximate likelihood-ratio test for ordinal response models. Stata Tech Bull. 1998; 42: 24-7.