♲ EPA

CatReg Software
User Manual
(R-Version)

CatReg Software
User Manual
(R-version)

National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

GLOSSARY OF TERMS

Akaike Information Criteria (AIC) - The deviance (-2 times the log of the maximized value of the likelihood function) + 2 times the number of parameters in the model.

Categorical regression - A model expressing the probabilities of different response categories as functions of explanatory variables.

Cluster sample - A data set comprised of subsamples of data from common sources. For example, a data set may contain several data records per laboratory from several different laboratories. The subgroup of data records from each individual laboratory would represent a cluster sample.

Cumulative odds regression - Ordinal regression model for directly modeling the probabilities of exceeding different severity levels.

Deviance - The minimized value of twice the negative logarithm of the likelihood function.

Extra risk concentration (ERC) - See Section 4.2.

ERC settings - Three numbers: ERC percentile, ERC severity level, and ERC percentile for confidence intervals (see Sections 5.1, and 8).

Filtering - Exclusion of selected data records from the analysis. This capability "filters out" selected data without altering the data input file.

Generalized estimating equation - An equation depending on the data and the parameter values, such that solving for the parameter values yields consistent estimates.

Hierarchical models - An ordered series of models, such that each model is a special case of the next one in the series.

Inf - Infinite value

Interval censored data - Data for which the response is known only to lie in an interval of values.

Likelihood function - For categorical response data, a model for the joint probability of the observed data values, expressed as a function of the model parameters.

Link function - A function applied to the categorical response probability to transform the categorical regression model to linear units.

Meta-analysis - The analysis of data from multiple studies to determine overall trends and increase power.

Na - Not available

NaN - Not a number

Ordinal data - Data reported as ordered categories. The order is meaningful, but the numerical difference between ordered categories is not.

Parallelism - The coefficients of concentration and time in the exposure-response model (i.e., the probability function) do not change with severity level (parallelism applies to Model 1, the cumulative odds model, but not to Model 2, the unrestricted cumulative model).

Probability function - The function of the explanatory variables that gives the probability of exceeding a given severity level.

Proportional odds regression - Ordinal regression in which the log-odds of exceeding different severity levels are parallel across severity categories. It is a special case of cumulative odds regression with the logistic link function.

Stratification - To create subsets of data by allowing the model parameters to vary by subset. Covariate information such as species, sex, and target organ may be used as a basis for creating the sub sets

1. INTRODUCTION

CatReg is a computer program developed to support toxicologists and health scientists in the conduct of exposure-response analyses, most often for controlled animal experiments. "Exposure" has two components: exposure level, indicated by a concentration or dose of the agent of interest, and exposure duration, when it varies within the data. "Response" refers to occurrence of a detrimental health effect of a user-defined level of severity. More specifically, effects observed in toxicological studies are assigned to ordinal severity categories and associated with the exposure conditions (e.g., concentration and duration) under which the effects occurred. "Ordinal" here means that the categories have a natural ordering in terms of severity or strength of response, but the spacing between ordinal scores is not subject to direct interpretation.

For example, response might have four levels of severity coded as: 0 = "no adverse effect", 1 = "mild adverse effect", 2 = "moderate/severe effect", 3 = "lethal effect". An ordinal response of 2 is higher than a response of 1, but the difference is not necessarily the same as the difference between 3 and 2. The simplest case is dichotomous response data, with just two severity levels, such as: 0 = "no adverse effect", 1 = "adverse effect". If data are reported on a continuous scale, such as mean and standard error of respiratory rate depression, the user can distribute the total number of experimental subjects over the severity levels using a method discussed in Section 3.2.3 and Appendix A.

CatReg provides two basic models, with variations to be explained, to relate the probabilities of the different severity categories to exposure level and exposure duration, taking user-defined covariates into account (e.g., species, gender, target organ, etc.). The parameters in the models are an intercept term and coefficients of concentration and duration, either of which may be log-transformed (to the base 10, denoted as "log," "log10," or "$\log_{10}$"). Model 1, the cumulative odds model, allows the intercept term to vary with severity level, but not the coefficients of concentration and duration. Model 2, the unrestricted cumulative model, allows any of the parameters to vary with severity level. The probability that a specified severity level or worse will occur increases as concentration or duration increases. The user can choose for either Model 1 or Model 2 to conform to the logistic, normal, or Gumbel cumulative probability distribution (see Section 4.1 and Appendix B). There is a function (called the link function) in each case that transforms the probability for each severity level to a linear function of the

unknown parameters, the format of a linear statistical model. The link functions are the logit, probit, and cloglog (complementary log-log) functions for the logistic, normal, and Gumbel cumulative probability distributions, respectively. The parameter estimates and their statistical characteristics, including standard errors and significance levels, are routinely output by CatReg, along with an analysis of deviance table to assess model fit and a table of estimates of extra risk concentrations (concentrations at which extra risk is a user-specified value) for 1, 4, 8, and 24 hour exposure durations.

CatReg was developed for, but is not limited to, meta-analysis of toxicology data. *Meta-analysis* refers to the analysis of data or results from multiple studies simultaneously. Meta-analysis becomes valuable when individual experiments are too narrow to address broad concerns. For example, in acute inhalation risk assessment, it is important to investigate the combined effects of concentration and duration of exposure, but few published experiments vary both the concentration and the duration of exposure (Guth et al., 1997). By combining information from multiple studies, the contribution of both concentration and duration to toxicity can be estimated. Moreover, the combined analysis allows the analyst to investigate variation among experiments, an important benchmark for the level of model uncertainty.

Different exposure-response experiments may consider the same or different toxicological endpoints, and toxicological judgment is required to determine if, and when, two different endpoints, or gradations of the same endpoint, are of comparable severity. A relatively simple example is analysis of mortality studies, with two severity levels: 0 = "not lethal", 1 = "lethal". The same endpoint is used for all studies and no intermediate degrees of health gradation are addressed. A little more complicated example might involve a single health effect, or mode of action, but with more than one severity level corresponding to manifestations of progressive "stages" of development.

Where studies address dissimilar endpoints that may be the consequence of different modes of action, particular care needs to be exercised to decide if comparable severity levels can be assigned across studies. It may not be reasonable to include all studies in the same analysis. For example, two toxicology experiments might report stages of anesthesia while another reports suppression of the shock-avoidance response. It might be the case that a toxicologist can confidently assign endpoints of the first two studies to comparable severity levels, but not be able to include the third study. In that case, one analysis could address the first two studies and a

6

second analysis the third study, since the studies cannot all be put on a "toxicologically equivalent" severity scale for analysis together.

A critical feature of CatReg is the capability to support data analysis needed for exposure-response modeling, that includes assessing and comparing how well models fit the data, testing for differences across studies and the significance of covariates within single or pooled studies, detecting outliers, etc. The program has a wide range of options to facilitate sensitivity analysis and produces numerous plots in addition to test results. In general, CatReg should be viewed as a statistical tool for developing an exposure-response curve and addressing related questions. A thorough analysis may require numerous executions of CatReg, ideally guided by both toxicological and statistical considerations.

The U.S. Environmental Protection Agency's (EPA's) National Center for Environmental Assessment encourages the broad application of this software. In this document, however, EPA has chosen to focus on the application of the software to the assessment of adverse effects associated with acute inhalation exposure. The description of data input files for the software reflect this application. The user is free to modify the input fields to support other applications. Appendix B provides additional technical description of the statistical methods used by the program. CatReg is designed to work with R, and the user must have access to this software to execute the CatReg program (http://cran.r-project.org/). Although familiarity with R may assist the user, an understanding of the R programming language is not required.

In the following text, `Courier New` font is used in most places for words that appear on the screen. For example, to plot confidence intervals for extra risk after estimation of the exposure-response curve, at the prompt (usually >) the user would input `confplot()`, that may include options within the parentheses.

2.   INSTALLING CatReg

R is required to run CatReg. These instructions apply to Windows 98/2000/XP operating systems. To obtain and install the source code for CatReg, follow the steps below.

(1).   Download the latest version of R, from the URL:http://cran.r-project.org/

(2).   Create a sub-directory (folder) named "CatReg" under r221, make it look used.) like:C:\Program Files\R\r221\CatReg.(Note: "r221" corresponds to the current R version number. This subdirectory name will change according to the version of R being used.)

(3). Download and save the attached five R source code files - catreg.R, ecdata.R, Util.R, Winplot.R, and Winstall.R to the "CatReg" folder.

(4).   Set the working directory in R to "CatReg". To do this, open R and enter the command:
        setwd("C:\\Program Files\\R\\r221\\CatReg")
        At the R command prompt, type getwd() to verify that the right working directory is set.

(5).   Select "File" from the main menu, and then "Source R code..." Go to the "CatReg" subdirectory, and double click "Winstall.R". This will automatically source all the five CatReg source files.

(6).   In R command prompt, type in "catreg()", and the CatReg program should be running.

Steps 4 -6 above must be performed each time CatReg is run in R. If you are a frequent CatReg user, you can add the following three R commends to the ".Rprofile" file in the "etc" subdirectory (e.g., using wordpad). By doing this, CatReg can be run without performing steps 4 -5.

```
setwd("C:/Program Files/R/r221/CatReg")
source ("C:/Program Files/R/r221/CatReg/Winstall.R")
load ("C:/Program Files/R/r221/CatReg/.RData")
```

3.   CREATING INPUT RECORDS

3.1.   Data Types

When the source document for an experiment does not report the outcome for individual subjects, or otherwise report the incidence of different health effects, the data may not be suitable for CatReg. For example, a report of "mild" pathology for a treatment group might mean that a few or many in the group manifested that response or that the "mild" response was the most common, with both lesser and more severe effects also present in the group. In either case, there is not sufficient information to divide a treatment group into incidence of severity categories.

It is sometimes reasonable to represent a health outcome measured on a continuous scale as categorical data. Continuous data from acute studies, such as enzyme activities, tidal volume, respiratory rate, blood pressure, etc., often are reported as a mean value, with a measure of dispersion, such as the standard error or standard deviation, for each treatment group. To convert these data to severity levels for CatReg, each severity level needs to be equated to an interval of

values on the continuous scale. For example, if the full range of responses is 0 to 100, the user might decide to classify outcomes 0 to 20 as "no effect", 21 to 40 as a "mild adverse effect", 41 to 65 a "moderate adverse effect", and 66 to 100 "severe effect". The mean for a treatment group falls into a single severity level, but some of the individual responses of subjects in the group may have been dispersed over adjacent severity levels.

Knowing the mean and standard deviation (or standard error that can be converted to a standard deviation by multiplying by the square root of the size of the treatment group) and assuming a distribution for the continuous data (e.g., normal), an estimate can be made of the incidence at each severity level (see Appendix A for details). The estimated incidence figures need not be whole numbers, but must still sum to the total group size. Incidence estimation is not possible if the mean is reported without a measure of dispersion.

3.2.   The Data Input File

The same category system of severity levels must be used for all data in an input file. Considerable toxicological judgment may be required for classification of various health effects into severity levels and for achieving comparability across experiments. When that can not be done for all the studies of interest, it may be necessary to group the studies into more than one input file. Classification judgments must be made systematically according to documented criteria.

The minimum number of severity levels is two (severity levels coded as 0 and 1, corresponding to absence or presence of an effect) and the maximum number is four. Suggested severity categories for a three-category classification are "no adverse effect", "adverse effect", and "lethal effect", coded as severity levels 0, 1, 2, respectively. A four-category scheme might be "no adverse effect", "mild adverse effect", "moderate/severe effect", and "lethal effect", coded as 0, 1, 2, 3, respectively.

In some toxicology studies, it may not be possible to score all response data completely. Consider a four-category scoring system in which 0 = "no observable effect," 1 = "mild effect," 2 = "moderate effect," and 3 = "severe effect." Published data from an animal mortality study may not include nonlethal outcomes; therefore, the response score for an animal that survives is uncertain, or "censored." That score is known to be less than 3, but it is not known whether the score should be 0, 1, or 2. Such an observation is said to be "interval censored." Another

situation where the response score may be interval censored is in combining data from experiments with different endpoints. For some endpoints, it may not be clear from the toxicology whether a specific response should be considered "mild" or "moderate." An interval censored analysis simply could report that the response is either 1 or 2, but the specific score is not known.

The ability to include partial information about the ordinal scores is one of the important features of CatReg. CatReg incorporates this type of partial information in an interval-censored analysis. In general, interval censoring occurs if the response is known only to lie in an interval of potential values. Such intervals are specified in CatReg by supplying the lower and upper limits of the known range for each observation. Suggested codes to indicate species and sex in an input file are provided in Table 3-1.

TABLE 3-1. RECOMMENDED CODES FOR SPECIES AND SEX

Species	Code	Sex	Code
Human	HU	Female	F
Rat	RT	Male	M
Mouse	MU	Both sexes	B
Rabbit	RB		
Guinea pig	GP		

Each column of the user input file is referred to as a data field, with the first record (row) being variable names and all subsequent rows containing data for the variables. CatReg requires information for four- six data fields, depending on the data. The names of these data fields and the corresponding default variable names that CatReg looks for in the user input file are shown in Table 3-2. For example, "conc" refers to a data field for exposure concentration and CatReg looks for the variable name "mg/m3" to identify that field, unless the variable name has been changed from the default (see Section 10.1). The default variable name for "conc" might be changed, for example, if the concentration used in experiments is different from milligrams per cubic meter.

Using the default variable names in Table 3-2 as an example, the user input file must include data in each record (beyond the first that contains variable names) for variables "mg/m3", "Nsub", "Incid", and "SevLo". Data are also required for the variable "Hours" unless

all exposure duration times are equal, in which case it can be omitted (CatReg uses 1 as the default value in that case). The variable SevHi is required if the severity level for one or more records is entered as a range, but not otherwise. For example, if the severity level for a record is

TABLE 3-2. DATA FIELDS THAT MAY BE REQUIRED

Data Field	Variable Name	Description
Conc	mg/m3	Exposure concentration. Use of the human equivalent concentration is recommended. Always required.
Time	Hours	Exposure duration in hours. Required unless all durations are equal. (Default value = 1.)
Nsub	Nsub	Number of subjects in a treatment group. Always required.
Incid	Incid	Incidence of severity level (or severity range) for the record. Always required.
Loscore	SevLo	Lowest severity level for the record. Always required.
Hiscore	SevHi	Highest severity level for the record. Required if response spans more than one severity level.

a range such as level 1 to level 2, then SevLo is entered as 1 and SevHi is entered as 2. Data records for which SevLo $\neq$ SevHi are referred as *censored* data. If only one severity level applies, e.g. level 1, then SevLo = SevHi = 1.

A separate record is entered for each severity level (or range of severity levels in the case of censored data) observed in a treatment group. For example, if the user determines three severity classifications for health effects, denoted as 0, 1, and 2, then the outcome for a treatment group is represented as an incidence for each severity level that is observed.   To illustrate, a treatment group of size 10 might result in 3 subjects being classified at severity level 0, 4 at severity level 1, and 3 at severity level 2, which would require three consecutive records (consecutive rows of data) in the input file. The records for the treatment group must not only be consecutive but their incidence (values of Incid) must sum to the treatment group size (Nsub).

A severity level with no observations need not be entered as a separate record. For example, if a treatment group of size 10 had 6 subjects classified at severity level 0 and 4 at

severity level 1, then only two records would be required to enter the data. The value of Nsub would be 10 in both records; the value of Incid would be 6 in one record and 4 in the other.

The user's input file, with data separated by commas, must be housed in the "Start in" directory for R. Microsoft Excel spreadsheets can be used to construct a data file, but the file must be saved as a comma-delimited file with a "csv" extension, rather than as an Excel file with an "xls" extension. Because CatReg assumes that data are separated by commas rather than by blank spaces, spaces are interpreted as characters and should be avoided unless intended to be part of the data. For instance, ", MU," is distinct from ",MU,". Because R is also case-sensitive, "mU," is different from "Mu,".

Variables in addition to those in Table 3-2 may be added to the input file at the user's discretion, either for use in execution of CatReg or to facilitate organizing and keeping track of the data. The user can refer to them in the same manner as variables in the required fields when using CatReg options. For example, one might want to add "strain" as a variable to distinguish between two strains of mice and have CatReg test whether their exposure-response curves are significantly different in some respect, or add "Ref.id" to record the source of the data, even if it is not used during the execution of CatReg.

Three varied examples of input files follow, described as experimental results for hypothethical chemicals named chemx, chemy, and chemz. The data were generated by simulation, except the data for what is being called chemy that were constructed from a few experiments on exposure of rodents to hydrogen sulfide. The input file for chemx is an example of four experiments, one each on the four combinations of species (RT and MU) and target organs (C and L). The input file for chemy has a more complicated structure and illustrates how a toxicologist might determine the severity levels. The input file for chemz is an illustration of converting a continuous response to severity categories for use in CatReg.

3.2.1. Input File: chemx.csv.

Table 3-3 displays the first part of the input file for chemx.csv. Four experiments were conducted under identical exposure conditions, each consisting of 10 observations at each combination of four concentrations (mg/m3) and four duration (Hours), for a total of 64 treatment groups. The concentrations are 1259, 1585, 2000, and 2512 mg/m3; the durations are 1.25, 1.6, 2.0, and 2.5 hours. There are three severity levels: no adverse effect (SevLo = 0), mild adverse effect (SevLo = 1), moderate/severe effect (SevLo = 2). Two experiments are on mice

(Species = MU) and two are on rats (Species = RT), with one of the two experiments on each species reporting effects on the central nervous system (Target = C) and the other reporting effects on the liver (Target = L). "Exp" denotes an experiment number, "Group" the treatment group within the experiment, "Nsub" the number of subjects in the treatment group, and "Incid" the incidence in the treatment group of the severity level (SevLo) shown in the record (row of data).

TABLE 3-3. PART OF THE INPUT FILE CHEMX.CSV.

Exp.	Group	Species	Target	mg/m3	Hours	SevLo	Nsub	Incid
1	1	MU	C	1259	1.25	0	10	10
1	2	MU	C	1259	1.6	0	10	9
1	2	MU	C	1259	1.6	1	10	1
1	3	MU	C	1259	2	0	10	4
1	3	MU	C	1259	2	1	10	6
1	4	MU	C	1259	2.5	0	10	1
1	4	MU	C	1259	2.5	1	10	9
1	5	MU	C	1585	1.25	0	10	7
1	5	MU	C	1585	1.25	1	10	3
1	6	MU	C	1585	1.6	0	10	3
1	6	MU	C	1585	1.6	1	10	7
1	7	MU	C	1585	2	0	10	1
1	7	MU	C	1585	2	1	10	7
1	7	MU	C	1585	2	2	10	2
1	8	MU	C	1585	2.5	1	10	4
1	8	MU	C	1585	2.5	2	10	6
1	9	MU	C	2000	1.25	0	10	7
1	9	MU	C	2000	1.25	1	10	3
1	10	MU	C	2000	1.6	0	10	2
1	10	MU	C	2000	1.6	1	10	5
1	10	MU	C	2000	1.6	2	10	3
1	11	MU	C	2000	2	1	10	6
1	11	MU	C	2000	2	2	10	4
1	12	MU	C	2000	2.5	1	10	4
1	12	MU	C	2000	2.5	2	10	6
1	13	MU	C	2512	1.25	1	10	9
1	13	MU	C	2512	1.25	2	10	1
1	14	MU	C	2512	1.6	1	10	5
1	14	MU	C	2512	1.6	2	10	5
1	15	MU	C	2512	2	1	10	2
1	15	MU	C	2512	2	2	10	8
1	16	MU	C	2512	2.5	2	10	10
2	1	RT	C	1259	1.25	0	10	10
2	2	RT	C	1259	1.6	0	10	9
2	2	RT	C	1259	1.6	1	10	1
2	3	RT	C	1259	2	0	10	9
2	3	RT	C	1259	2	1	10	1

The variable names mg/m3, SevLo, Nsub, and Incid are required and an error message is printed if any of them is missing. In this example, the exposure durations vary so the variable Hours is included. SevHi would have been included as a variable if the severity level had been censored for one or more records (spanned more than one severity level). For example, to make the 10 subjects in the first record (row of data) classified as severity level 0-1, the variable SevHi would be added to the input file and the first record would remain unchanged except for SevHi = 1. In that case, subsequent records that are not censored would be given the same value for SevLo and SevHi. For example, the second record indicates that 9 subjects were classified at severity level 0. If SevHi were included as a variable, then SevHi would be set to 9 for that record, making SevLo = SevHi = 9.

It may be noted that variables Exp., Group, Species, and Target are names created by the user. Species and Target were included in this case to be able to distinguish between species and target organ in the data analysis, but other names could be used in their place. Exp. and Group were added by the user to facilitate record keeping. The variable Group is not required by CatReg but records for the same treatment group must be together in the input file, all with the common value of Nsub and values of Incid that sum to Nsub. Adding a variable such as Group provides a convenient check of the data for the user. CatReg determines treatment groups by reading records until the values of Incid sum to Nsub, then starting over with the next record.

3.2.2. Input File: chemy.csv.

Table 3-4 is part of a larger input file that was constructed for exposure of rodents to hydrogen sulfide. Only part of the available experimental data are used here for illustration, so it is referred to as "chemy" instead of hydrogen sulfide. It illustrates a more elaborate coding system and some other features not included in the preceding example, e.g. censoring, and provides a realistic example for discussion of toxicological judgment in severity classification. The available studies varied on the organ sites and endpoints examined and a four-category system of severity levels was implemented: no adverse effect (0), mild adverse effect (1), moderate/severe effect (2), and lethal effect (3). Again, CatReg expects to find variables mg/m3, Nsub, Incid, and SevLo, at a minimum, and Hours as well, if exposure duration varies, and SevHi if there are any censored data. Also (again) notice that a separate record is required for

each different severity (or range of severity in the case of censored data) in a treatment group, and that Incid sums to Nsub for records in the same treatment group.

"Ref.id" (reference identification) is a number assigned by the user to the source of the information in the record. In the example, Ref.id = 20938 is the source of the material used to construct the first 10 data records shown. "Exp." identifies experiments within each Ref.id, numbered sequentially from 1; "Group" numbers treatment groups within each experiment (i.e., subjects alike with respect to all methods and materials variables); and "Nsub" is the number of subjects in each group. A separate record (row) is entered for each severity level (or range of

TABLE 3-4. PART OF INPUT FILE CHEMY.CSV

Marker	Ref.id	Exp	Group	Species	Sex	mg/m3	Hours	Target	Endpoint	Nsub	Incid	BestNum	SevLo	SevHi	Censored
1	20938	1	1	RT	B	330	6	Resp	Lethality	26	26	1	0	2	y
2	20938	1	2	RT	B	390	6	Resp	Lethality	26	20	2	1	2	y
3	20938	1	2	RT	B	390	6	Resp	Lethality	26	6	3	3	3	n
4	20938	1	3	RT	B	460	6	Resp	Lethality	26	23	3	3	3	n
5	20938	1	3	RT	B	460	6	Resp	Lethality	26	3	2	2	2	n
6	20938	2	1	MU	B	360	6	Resp	Lethality	26	23	2	1	2	y
7	20938	2	1	MU	B	360	6	Resp	Lethality	26	3	3	3	3	n
8	20938	2	2	MU	B	420	6	Resp	Lethality	26	13	3	3	3	n
9	20938	2	2	MU	B	420	6	Resp	Lethality	26	13	2	1	2	y
10	20938	2	3	MU	B	500	6	Resp	Lethality	26	26	3	3	3	n
11	61831	1	1	RT	M	14	4	Resp	N lavage	12	12	0	0	0	n
12	61831	1	2	RT	M	278	4	Resp	N lavage	12	12	1	0	1	y
13	61831	1	3	RT	M	556	4	Resp	N lavage	12	12	2	1	2	y

Notes:

Marker - Record number.
Ref. id - Source identifier.
Exp - Experiment number within a source.
Group - Treatment group number (within an experiment).
Species - Species.
Sex - Sex.
mg/m3 - Exposure concentration.
Hours - Exposure duration.

Target - Target organ.
Endpoint - Toxic endpoint.
Nsub - Number of subjects in treatment group.
Incid - Number of animals responding.
BestNum - Analyst's best estimate of severity category for censored data; same as SevLo and SevHi for noncensored data.
SevLo - Lowest applicable severity level.
SevHi - Highest applicable severity level.
Censored - Severity level reported as a range if "y".

severity levels) in a treatment group. "Marker" just numbers the data records. The severity levels are entered under "SevLo", the lowest possible severity level for the record, and "SevHi", the highest severity level for the record.

When SevLo ≠ SevHi, "y" is entered under "Censored"; "n" when SevLo = SevHi. In the example (Table 3-4), the variable Censored has been added by the user to readily distinguish between records with a single severity level and those with a range of severity levels; it is not required by CatReg. The variable BestNum also has been added by the user to indicate the most likely severity level when SevLo ≠ SevHi; otherwise, the common value for SevLo and SevHi is entered for BestNum. To use the scores in the BestNum column, the default variable names SevLo and SevHi must both be changed to BestNum, as described in Section 10.1.

The first reference (Ref.id = 20938) reported two experiments, one with mice (Species = MU) and one with rats (Species = RT), with both sexes in both experiments (Sex = B). There is a user-defined variable coded to indicate the target organ (e.g., Target = Resp) and the primary endpoint (Endpoint = Lethality), both of which were the same for both experiments. The first group (Group = 1) under the first experiment (Exp = 1) is for rats (Species = RT) of both sexes (Sex = B) exposed at a concentration of 330 mg/m³ (mg/m3 = 330) for 6 h (Hours = 6). There were 26 rats at risk (Nsub = 26). The user-defined severity classification was the same for all 26: severity level 0 to 2 (SevLo = 0, SevHi = 2), with severity level 1 the best guess for a single severity level (BestNum = 1). Because the results variables were the same for all 26 subjects, only one record is needed to record the data for the whole treatment group.

The sixth and seventh records (Markers 6 and 7) are for the first treatment group (Group = 1) of the second experiment (Exp = 2) in Ref.id = 20938. Mice (Species = MU) of both sexes (Sex = B) were exposed to 360 mg/m³ (mg/m3 = 360) for 6 h (Hours = 6). Two records are required because there are two distinct severity classifications: one for 23 subjects (Incid = 23) with severity level 1 to 2 (SevLo = 1, SevHi = 2) and severity level 2 the best guess (BestNum = 2) and the other with three subjects at severity level 3 (SevLo = 3, SevHi = 3) and BestNum = 3. Six records in the example have different values for SevLo and SevHi, as indicated by Censored = y.

In Ref.id = 20938 of Table 3-4, the effect severity for concentrations at which no subjects died were censored 0 to 2 (e.g., Marker 1) for no adverse effects to severe adverse effects, because the effects were unknown. Survivors from groups in which some subjects died (e.g., Marker 2) were assumed to have suffered adverse effects and were censored 1 to 2 because the

17

effects could have been mild to severe. Survivors from groups in which most of the subjects died (e.g., Marker 5) were assumed to have suffered severe effects. For the 360-, 390-, 420-, and 460-mg/m^3 H$_2$S exposures, there is one record for the number of subjects exhibiting lethal effects, and another for the number of subjects exhibiting nonlethal effects. Only one record was made for exposure to 330 mg/m^3, because all subjects were assumed to exhibit effects of severity 0 to 2, and one record was made for exposure to 500 mg/m^3, because all subjects died (severity 3).

Ref.id = 61831 of the example illustrates a case where health effects are described for each treatment group, for which the severity levels can be decided, but incidence of the severity levels is unknown. These data are not suitable for CatReg because correct values for Incid are unknown. Nevertheless, one might want to test the sensitivity of the results of CatReg to different assumptions regarding the incidence values and it provides a useful example for discussion of deciding severity levels. It is included here primarily as an example of deciding severity levels from toxicological data. The data are from a study by Lopez et al. (1987), in which groups of 12 rats were exposed to 0, 14, 278, and 556 mg/m^3 H$_2$S for 4 h. Incid = 12 is used in the example for illustration, which is equivalent to assuming that all 12 rats in each treatment group have the same severity classification shown for the group as a whole.

Data from the Lopez et al. study were assigned to severity levels as follows. Groups of four rats were killed at 1, 20, and 44 h after 4-h exposure for the examination of biochemical indicators of injury and inflammatory response in the respiratory tract. Nasal lavage fluid was examined for lactate dehydrogenase (LDH), alkaline phosphatase (ALP), protein, and number of nucleated cells. Bronchoalveolar lavage (BAL) fluid was examined for activities of LDH, ALP, and γ-glutamyl transpeptidase. All measurements were reported as means, plus or minus standard deviations. No changes in any parameters were noted among rats exposed to 14 mg/m^3 H$_2$S. The only parameter significantly different from controls in nasal lavage fluid at all post-exposure time periods was increased cellularity, which was significant at the 556-mg/m^3 exposure. In BAL fluid, LDH activity was elevated at 44 h post-exposure, and ALP activity was significantly decreased at 20 and 44 h after the 278-mg/m^3 exposure. At all post-exposure durations for the 556-mg/m^3 exposure, protein concentration and LDH activity were elevated, but ALP was decreased.

Because a number of biochemical and cellular parameters were measured at several post-exposure periods, significant changes were considered to be adverse only if the changes were

18

still significant at the last post-exposure measurement. In other words, reversible changes were classified as no-observed-adverse effects. Table 3-5 shows how these effects were categorized using the four-category severity scheme. Because no changes were noted after exposure to 14mg/m^3 H$_2$S, effects were coded 0 for no adverse effect. Effects at the 278-mg/m^3 H$_2$S exposure were estimated to range from severity 0 to 1 (no adverse effect to mild adverse effect) because significant changes in two biochemical parameters were observed. The adversity of those changes was uncertain but assumed to be less than severe. Effects caused by exposure to 556 mg/m^3 H$_2$S were estimated to range from severity 1 to 2 (mild to severe adverse effect) because of changes in the activities of several enzymes and nasal cytopathology. No deaths occurred during this study, so category 3 (lethality) was not used.

Table 3-5. EXAMPLE OF SEVERITY CATEGORIZATION FOR NONLETHAL EFFECTS

Exposure Concentration (mg/m^3 H$_2$S)	Statistically Significant Effects Reported	Severity Score	Censored
14	None	0	No
2748	↑LDH, ↓ALP	0-1	Yes
556	↑Protein, ↑LDH, ↓ALP, ↑cellularity, nasal cytopathology	1-2	Yes

3.2.3. Input File: chemz.csv

An artificial example of how continuous data might be coded is displayed in Table 3-6. It is assumed that the standard error was included along with the mean for each treatment group. In this case, the incidence of each severity in a group was estimated from the mean and standard error, which produces fractional subjects. The method to estimate incidence at each severity from group means and a measure of variability is described in Appendix A.

Male rats in treatment groups of 10 each were exposed to a toxicant at various concentrations and durations. Adverse effects occurred in the respiratory tract, with severity indicated by lung weight. For illustration, a four-category severity classification is used, with lung weight 0 to 20 classified as "no effect", 21 to 40 as "mild adverse effect", 41 to 65 as "moderate adverse effect", and 66 to 100 as "severe effect". The mean for a treatment group

TABLE 3-6. PART OF INPUT FILE CHEMZ.CSV

Marker	Ref.id	Exp	Group	Species	Sex	mg/m3	Hours	Target	Endpoint	Nsub	Incid	BestNum	SevLo	SevHi	Censored
1	1	1	1	RT	M	330	2	Resp	Lung wt	10	1	0	0	0	n
2	1	1	2	RT	M	360	2	Resp	Lung wt	10	1	1	1	1	n
3	1	1	3	RT	M	390	2	Resp	Lung wt	10	1	1	1	2	y
4	1	1	4	RT	M	410	2	Resp	Lung wt	10	1	2	2	2	n
5	1	1	5	RT	M	460	2	Resp	Lung wt	10	1	2	2	2	n
6	1	2	1	RT	M	460	1	Resp	Lung wt	10	1	1	1	1	n
7	1	2	2	RT	M	510	1	Resp	Lung wt	10	1	1	1	1	n
8	1	2	3	RT	M	560	1	Resp	Lung wt	10	1	2	2	2	n
9	1	2	4	RT	M	610	1	Resp	Lung wt	10	1	2	2	2	n
10	1	3	1	RT	M	560	0.5	Resp	Lung wt	10	1	1	1	1	n
11	1	3	2	RT	M	610	0.5	Resp	Lung wt	10	1	1	1	2	y
12	1	3	3	RT	M	660	0.5	Resp	Lung wt	10	1	2	2	2	n
13	1	3	4	RT	M	710	0.5	Resp	Lung wt	10	1	2	2	2	n
14	2	1	1	RT	M	330	2	Resp	Lung wt	10	7.1	0	0	0	n
15	2	1	1	RT	M	330	2	Resp	Lung wt	10	2.9	0	1	1	n
16	2	1	2	RT	M	360	2	Resp	Lung wt	10	2.9	1	0	0	n
17	2	1	2	RT	M	360	2	Resp	Lung wt	10	7.1	1	1	1	n
18	2	1	3	RT	M	390	2	Resp	Lung wt	10	4.9	1	1	1	n
19	2	1	3	RT	M	390	2	Resp	Lung wt	10	5.1	1	2	2	n
20	2	1	4	RT	M	410	2	Resp	Lung wt	10	1.2	1	1	1	n
21	2	1	4	RT	M	410	2	Resp	Lung wt	10	8.8	2	2	2	n

TABLE 3-6 (CONT'D). PART OF INPUT FILE CHEMZ.CSV

Marker	Ref.id	Exp	Group	Species	Sex	mg/m3	Hours	Target	Endpoint	Nsub	Incid	BestNum	SevLo	SevHi	Censored
22	2	1	5	RT	M	460	2	Resp	Lung wt	10	1.5	1	1	1	n
23	2	1	5	RT	M	460	2	Resp	Lung wt	10	7.2	2	2	2	n
24	2	2	1	RT	M	460	1	Resp	Lung wt	10	2.9	1	0	0	n
25	2	2	1	RT	M	460	1	Resp	Lung wt	10	7.1	1	1	1	n
26	2	2	2	RT	M	510	1	Resp	Lung wt	10	1.3	0	0	0	n
27	2	2	2	RT	M	510	1	Resp	Lung wt	10	7.9	1	1	1	n
28	2	2	2	RT	M	510	1	Resp	Lung wt	10	0.7	2	2	2	n
29	2	2	3	RT	M	560	1	Resp	Lung wt	10	3.5	1	1	1	n
30	2	2	3	RT	M	560	1	Resp	Lung wt	10	6.5	2	2	2	n
31	2	2	4	RT	M	610	1	Resp	Lung wt	10	1.6	1	1	1	n
32	2	2	4	RT	M	610	1	Tesp	Lung wt	10	8.1	2	2	2	n
33	2	3	1	RT	M	560	0.5	Resp	Lung wt	10	9.4	1	1	1	n
34	2	3	2	RT	M	610	0.5	Resp	Lung wt	10	4.9	1	1	1	n
35	2	3	2	RT	M	610	0.5	Resp	Lung wt	10	5.0	2	2	2	n
36	2	3	3	RT	M	660	0.5	Resp	Lung wt	10	3.6	1	1	1	n
37	2	3	3	RT	M	660	0.5	Resp	Lung wt	10	6.5	2	2	2	n
38	2	3	4	RT	M	710	0.5	Resp	Lung wt	10	1.6	1	1	1	n
39	2	3	4	RT	M	710	0.5	Resp	Lung wt	10	8.1	2	2	2	N

falls into a single severity level, but some of the individual responses of subjects in the group may have been dispersed over adjacent severity levels.

4.   TECHNICAL BACKGROUND: MODELS AND EXTRA RISK

4.1.   Exposure-response Models

Let Y denote a dependent variable that represents the severity or intensity of the response. Assume Y is an ordinal score taking one of the values $(0,1,...,S)$. A score of 0 corresponds to the lowest severity (e.g., no adverse effect), and a score of S corresponds to the highest severity (e.g., lethal, in a toxicological context). Categorical regression is a method for modeling the probability distribution of Y as a function of the explanatory variables, concentration (C) and duration (T). It employs a generalized linear model (McCullagh and Nelder, 1989) for the dependence of the probabilities of different severity categories on the explanatory variables.

CatReg provides a choice of two models, Model 1 (the cumulative odds model) described by Eq. 4-1a and Model 2 (the unrestricted cumulative model) described by Eq. 4-1b, with variations on both as described below. CatReg refers to any model of the form of Model 1 as a cumulative odds model because the model is expressed in terms of the cumulative probabilities, or odds, for $Y \geq s$. Note that Model 1 is a special case of Model 2 wherein parameters β_{1s} and β_{2s} do not depend on s (which denotes severity level, as discussed below). A primary use of fitting Model 2 is to test whether the simpler Model 1 is adequate. Model 2, although more general than Model 1, has the undesirable feature that the regression lines for different severity levels may cross. Often the crossing is well outside the range of values of interest, so the model can be used to make empirical risk estimates. The user has the option to add an additional parameter, γ, which represents a hypothetical background concentration, in some circumstances (see Section 5.4). For $s = 1,2,\ldots,S,$

Model 1

$$\Pr(Y \geq s | C,T) = H[\alpha_s + \beta_1 * f_1(C) + \beta_2 * f_2(T)]$$ **Eq. 4-1a**

Model 2

$$\Pr(Y \geq s | C,T) = H[\alpha_s + \beta_{1s} * f_1(C) + \beta_{2s} * f_2(T)]$$ **Eq. 4-1b**

The left-hand side is read as follows: the probability that a response of severity level *s* or greater occurs, given that concentration is *C* and time is *T* (time refers to exposure duration). No expression for *s* = 0 is included because this is the minimal category, and *Y* is always greater than or equal to 0 (i.e., $\Pr(Y \geq 0|C,T) = 1$). The right-hand side is described as follows:

- *H* is a probability function taking values between 0 and 1, for which the user has three choices: (1) logistic, (2) normal, and (3) Gumbel (described further in Appendix B).

- The parameter α_s is the intercept for severity level s, $s = 2,...,S$ (to be called the intercept or severity parameters). The severity parameters are ordered as $\alpha_1 \geq \alpha_2 \geq ... \geq \alpha_S$. This constraint is a consequence of the requirement that the probability of exceeding a lower score is larger than the probability of exceeding a higher score for any fixed levels of *C* and *T*.

- In Model 1, the parameter β_1 determines the dependence of the response on concentration (C), whereas β_2 determines the dependence on time (T). In Model 2, the parameters are also indexed by s because they may change values with severity level s.

- Current choices for f_1 and f_2 are "untransformed " and "base-10 logarithm ". Other transformations of *C* and *T* may be obtained by transforming the input data.

- Parameters are α_s, β_{1s} (to be called the coefficient of concentration), β_{2s} (to be called the coefficient of time or duration), for severity levels *s* running from 1 to *S*. All parameters may be stratified on variables, as discussed under "Stratifying" in Section 5.2.

The normal and logistic distributions are symmetric, each having median equal to zero. The Gumbel distribution is skewed, with a lower tail similar to that of the logistic distribution and a lighter upper tail. Figure 4-1 displays these three probability distributions. To compare the shapes effectively, the distributions have been rescaled to have medians = 0 and equal 25th percentiles (labeled as "EC25" on the horizontal axis). The scaled logistic and normal distributions are very close for much of the range and differ substantively only in the extreme tails. The Gumbel distribution is skewed, with a heavier tail on the left and a lighter tail on the right.

For each of the three choices of the probability function *H* there is an inverse function of *H*, called the link function, that transforms it to a simple linear function in concentration and duration. CatReg requests the name of the link function instead of the name of the probability function. The corresponding link functions (in parentheses) are: logistic (logit), normal (probit),
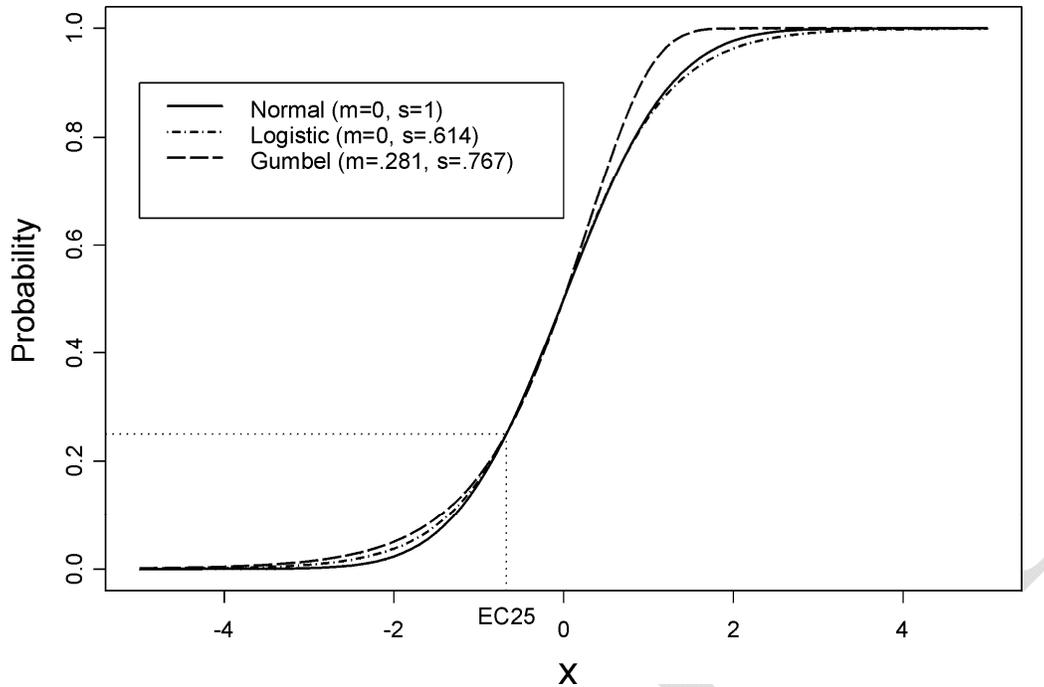
Figure 4-1. Normal, logistic, and Gumbel probability functions shifted and scaled to have equal medians and 25th percentiles.

and Gumbel (cloglog). There is further discussion of linking in Section 5.2 and Appendix B, that includes an example of how link functions may be derived from a basic assumption that the ordinal severity score corresponds to exceeding an underlying toxic response threshold.

4.2.   Extra Risk Concentration (ERC)

Extra risk at concentration C=c and time T=t, at severity level s, is defined as

$$\frac{\Pr(Y \geq s | C = c, T = t) - \Pr(Y \geq s | C = 0, T = t)}{1 - \Pr(Y \geq s | C = 0, T = t)}$$

Eq. 4-2

For q between 1 and 100, inclusive, ERCq, at time T = t, is the concentration c for which equation (4-2) equals q/100. For example, ERC10 at T = 2 (exposure duration of 2 hours) for severity level 1 is the value of c that satisfies

24

$$\frac{\Pr(Y \geq 1 | C = c, T = 2) - \Pr(Y \geq 1 | C = 0, T = 2)}{1 - \Pr(Y \geq 1 | C = 0, T = 2)} = 0.1 \qquad \textbf{Eq. 4-3}$$

In words, ERC10 at T = 2 for severity level s is the exposure concentration at which the probability is 0.10 of an adverse effect of level s or higher *due to exposure of two hours*, i.e., given the adverse effect would not have occurred from other causes ("background causes") during that time.

5.   RUNNING CatReg

This section will describe how to run CatReg after an input data file has been constructed, following discussion of some background material on the options for exposure-response curves available to the user.

5.1.   The Order of CatReg-User Interactions

It is assumed that an input file has been constructed in the format described in Section 3 and saved in the "Start in" directory for R, the same location of the CatReg script files, as described in Section 2. After launching R, open the commands window via the window menu on the toolbar, and enter `catreg()` at the R prompt (usually ">") followed by <CR> (press the "enter" key).   User interactions with CatReg will be described for a simple example of data from a single experiment. Input file chemx.csv is used with data from three of the four experiments filtered (i.e., ignored but not removed from the input file), leaving the single experiment on rats (RT) with response on the central nervous system (C). What appears on the monitor during execution of CatReg for this example is displayed in Table 5-1. Underscored section headings in boldface, e.g., **###(1)Begin setup ###**, have been added to facilitate explanation. Otherwise, boldface indicates user input from the keyboard.

The first section of Table 5-1, starting at **###(1)Begin setup ###**, consists of user queries to tell CatReg what options to use for an exposure-response curve. Setup options, discussed in the next section, refer to filtering, clustering, stratifying, choosing the scales for concentration and duration; choosing the exposure-response function (presented as a choice of the associated link function); specifying whether the exposure-response function should be formulated for the cumulative odds model or the unrestricted odds model (listed as a choice of

TABLE 5-1. MONITOR DISPLAY DURING EXECUTION OF CATREG RUN. INPUT FILE: CHEMX.CSV. FILTERED DATA: SPECIES =MU, TARGET = L. MODEL 2. LOGIT. SCALES: LOG10. ASSUMED ZERO BACKGROUND RISK.

```
> catreg()

CatReg - R Beta Version 1.1 (12/07/05)

Name of data file? chemx.csv
Read 1179 items
```

(1)Begin setup

```
 Variables:
 Exp. Group Species Target mg/m3 Hours SevLo Nsub Incid GpSize

 Filter the data on which variables? (none):
1: Species
2: Target
3:
Read 2 items

 Categories of Species:
 MU RT

 Categories to exclude? (none):
1: MU
2:
Read 1 item

 Categories of Target:
 C L

 Categories to exclude? (none):
1: L
2:
Read 1 item

 Variables:
 Exp. Group Species Target mg/m3 Hours SevLo Nsub Incid GpSize

 Define clusters using which variable(s)? (none):
1:
Read 0 items

 Variables:
 Exp. Group Species Target mg/m3 Hours SevLo Nsub Incid GpSize

 Stratify intercepts on which variables? (none):
1:
Read 0 items

 Variables:
 Exp. Group Species Target mg/m3 Hours SevLo Nsub Incid GpSize

 Stratify concentration (mg/m3) on which variables? (none):
1:
Read 0 items

 Variables:
```

Exp. Group Species Target mg/m3 Hours SevLo Nsub Incid GpSize

 Stratify time (Hours) on which variables? (none):
1:
Read 0 items

Log or linear scale for mg/m3?
Choices:

1: mg/m3
2: log10(mg/m3)

Selection: **2**

Log or linear scale for Hours?
Choices:

1: Hours
2: log10(Hours)

Selection: **2**

Link function?

1: logit
2: probit
3: cloglog

Selection: **1**

Model?

1: cumulative odds model
2: unrestricted cumulative model

Selection: **2**

Assume zero background risk
(i.e. response cannot occur at zero concentration)? (y): **y**

Name of the output? (chemx):

>>>> Output file is chemx.out

Stratification complete...

Continue? (y):

(2)Begin fitting exposure-response curve

Optimizing...

Computing estimated covariance matrix...

(3)Display and output summary results

Printout covariance matrix? (n): **y**

...The covariance matrix printed out in output file

Input file   : chemx.csv
Filtered data: Species Target : MU L
Model   : unrestricted cumulative model
Link   : logit

```
Clustering    : none
Message       :
Iterations    : 43 27
Deviance      : 167.6454
Residual DF   : 26
AIC           : 179.6454

Scale:
     Concentration: log10( mg/m3 )
     Duration     : log10( Hours )

Stratification:
     No Stratification on Intercept, Concentration and Duration.

Coefficients:
                 Estimate Std. Error   Z-Test=0     p-value
SEV1             -76.52406  12.431120  -6.155846    1e-05
SEV2            -104.22063  20.225024  -5.153054    1e-05
LG10CONC:SEV1    22.80200   3.719204   6.130882    1e-05
LG10TIME:SEV1    14.49540   3.002936   4.827075    1e-05
LG10CONC:SEV2    29.30517   5.806197   5.047222    1e-05
LG10TIME:SEV2    21.09271   4.453506   4.736202    1e-05

Writing summary data to file: chemx.out
```

(4)Output analysis of deviance table

Stratification complete...

(5)Calculate and output extra risk

Optimizing...

Calculating extra risk concentrations...

Percentile for ERC? (default=10):

Severity level for ERC? (default = 1):

Percentile for 1-sided upper and lower confidence intervals? (default=95):

ERC data computed...

Write ERC data to text files? (n):**y**

Writing ERC data to ascii files...

Done - see chemx.out for file names.

(6)Issue commands

```
>>>> Type `prplot(time=x)' or `prplot(conc=x)' to
>>>> graph probability with time or conc fixed at x.
>>>>
>>>> Type `catplot()' to display the ERC line and
>>>> the confidence lines for a particular stratum.
>>>>
>>>> Type `stratplot()' to display the ERC lines
>>>> for all strata.
>>>>
>>>> Type `confplot(10)' to display ERC confidence
>>>> intervals for duration=10.
>>>>
>>>> Type 'dataplot()' to plot points on duration-versus-concentration
```

```
>>>> axes without a response probability curve.
>>>>
>>>> Type 'devplot()' to provide a diagnostic plot of generalized deviance
>>>> residuals versus observation number, concentration, or duration.
>>>>
>>>> To remove previous plots type `rmplots()'.
>>>>
>>>> To change percentile or severity type `ecdata()'.

Warning message:
NAs introduced by coercion
```

models); whether to assume background risk is zero (an option when log-transform concentration has been selected) and there is no stratification; and naming the output file. In many of the CatReg queries, a default option, implemented by <CR> is given in parentheses. Otherwise, the user responds with a number from the selection of options presented, or with the name of a variable, which will be clear from the nature of the query.

To assist the user, when a variable name is needed CatReg prints the names of all variables. The user must respond with the variable name as shown, including the same use of lowercase and uppercase letters, as CatReg is case-sensitive. CatReg does not require the full name but just enough for it to be distinguished from the other variables. For "Species," only the "S" would be required if there is no other variable that begins with "S," but if "Sex" and "Sev" are also variables, then "Sp" would be required to uniquely identify "Species". For all queries for which multiple responses are possible, entries can be listed all on one line separated by a space, or entries can be typed on separate lines. <CR> on a line with no other user input moves the program on to the next query. Entries that are redundant or are unmatched name fragments will provoke an error message, and the query will be repeated. To abort the analysis at any step, enter "CTRL-C" for version 3.1 of R or "ESC" for higher versions.

The options and the user responses, shown in boldface in Table 5-1, are described as follows. The user keyed in Species and then Target as variables on which to filter (ignore) the data. CatReg then asked which categories of Species to exclude, to which the user keyed in MU to tell CatReg to ignore all records on the input file where Species = MU (i.e., where the value of the variable "Species" is "MU"). Similarly, CatReg was told to ignore all records in which Target =L. This left only the data records with Species = RT and Target = C. The user responses to the remaining queries were either the default or the number of the selection shown in boldface. The meaning of the options will be discussed subsequently.

For this example, Table 5-1 indicates that there was no clustering, no stratification of any of the parameters (intercept, coefficient of concentration (mg/m3), or coefficient of time (Hours)), concentration and time were both log-transformed, the logit link and unrestricted cumulative model were selected. The user elected to assume zero background risk (i.e., zero probability of severity level 1 or above when concentration is zero), an option that appears when concentration is log-transformed and there is no stratification; otherwise a parameter γ (gamma) would have been added to the model as a hyptothetical background concentration (see discussion of scaling in Section 5.2). The user chose the default name for the output file, "chemx", the name of the input file without the file type "csv"; the file type of the output file is "out".

CatReg then fit an exposure-response curve to the data by maximum likelihood and displayed the messages shown in the second section of Table 5-1 at **###(2)Begin fitting exposure-response curve ###**. At the section beginning with **###(3)Display and output summary results ###** , CatReg queried the user on whether to include the covariance matrix of the parameter estimates in the output file, wrote the output file, and displayed a summary portion of the output file on the monitor referred to as the *CatReg summary* (because the command catreg.summary()will also display it on the monitor; see Section 9.3). The CatReg summary is from the line Input file ... through the table of coefficient estimates. The table of coefficient estimates in the summary portion lists the parameters in the exposure-response curve, and the estimate, standard error, and significance level of each for a hypothesis that the parameter is zero. The correspondence between the parameters listed under Coefficients in Table 5-1 and those in Equation 4-1b are: SEV1 (α_1), SEV2 (α_2), LG10CONC:SEV1 (β_{11}), LG10CONC:SEV2 (β_{12}), LG10TIME:SEV1 (β_{21}), LG10TIME:SEV2 (β_{22}).

At **###(4)Output analysis of deviance table ###** in Table 5-1, CatReg calculated an analysis of deviance table and wrote it to the output file. The section starting with **###(5)Calculate and output extra risk ###** queried the user on options for the calculation of extra risk (ERC percentile, ERC severity level, and ERC percentile for confidence intervals, referred to collectively as the *ERC settings*) and whether to write the results to text files. The default responses were chosen in the example of Table 5-1 (ERC percentile = 10, ERC severity level = 1, ERC percentile for confidence interval percentile = 95). CatReg then calculated concentrations that will give extra risk of 0.10 for 20 exposure durations, equally

spaced and including the shortest and longest duration in the input file, applicable to severity level 1 with 95% one-sided upper and lower confidence intervals, and wrote them to a text file. The file was given the same name as the output file with the number "1" appended, i.e., "chemx1". If there had been stratification, a separate text file would have been written for each stratum and the names would have been "chemx1", "chemx2", … depending on the number of strata. Additionally, the ERC estimates and confidence intervals for exposure durations of 1, 4, 8, and 24 hours were automatically calculated and written to the output file.

The final section of Table 5-1, starting with **`###(6)Issue commands ###`**, enters the phase where the user determines what CatReg does by issuing commands. Commands have the following general objectives: testing hypotheses about exposure-response parameters, testing whether Model 2 can be reduced to Model 1, making graphs (the plotting options), and providing utility operations. The instructions for several commands were displayed on the screen, but no commands were executed.

The output file from the example shown in Table 5-1 (chemx, file type "out") is displayed in Table 5-2 and the ERC data that were written to a text file (chemx1, file type "txt") are shown in Table 5-3. Table 5-2 provides summary information such as the name of the input file, the setup options used, the table of coefficient estimates, an analysis of deviance table, and extra risk concentrations (ERCs) with upper and lower confidence bounds at exposure durations of 1, 4, 8, and 24 hours. Table 5-3 displays the ERC data written to a text file as requested by the user. It gives the ERC estimates, with confidence intervals, at the ERC percentile (10), severity level (1), and percentile for confidence intervals (95) that were set during the execution of CatReg (all are default values, in the section **`### (5)Calculate and output extra risk ###`** of Table 5-1).

5.2.   Setup Options—Discussion

Filtering. This option allows selected records to be removed (filtered) from the analysis, without removing them from the input file. It is used to fit the exposure-response curve to a subset of the data. One reason to filter data is to investigate how the fit changes when certain observations are excluded. For instance, a particular study may be suspect, and it may be desirable to compare the parameter estimates with and without the suspect study.

TABLE 5-2. OUTPUT FILE FROM EXAMPLE IN TABLE 5-1.

```
CatReg - R Beta Version 1.1 (12/07/05)

Source data file: chemx.csv

Type of analysis: Censored

Input file   : chemx.csv
Filtered data: Species Target : MU  L
Model        : unrestricted cumulative model
Link         : logit
Clustering   : none
Message      :
Iterations   : 43 27
Deviance     : 167.6454
Residual DF  : 26
AIC          : 179.6454

Scale:
     Concentration: log10( mg/m3 )
     Duration     : log10( Hours )

Stratification:
     No Stratification on Intercept, Concentration and Duration.

Coefficients:
               Estimate Std. Error   Z-Test=0     p-value
SEV1           -76.52406  12.431120  -6.155846     1e-05
SEV2          -104.22063  20.225024  -5.153054     1e-05
LG10CONC:SEV1   22.80200   3.719204   6.130882     1e-05
LG10TIME:SEV1   14.49540   3.002936   4.827075     1e-05
LG10CONC:SEV2   29.30517   5.806197   5.047222     1e-05
LG10TIME:SEV2   21.09271   4.453506   4.736202     1e-05

Variance:
                    SEV1         SEV2  LG10CONC:SEV1 LG10TIME:SEV1
LG10CONC:SEV2
SEV1          154.532745    1.3920254  -46.17931915   -25.7208267     -
0.5190940
SEV2            1.392025  409.0516115   -0.41877765     0.7452832     -
117.2769304
LG10CONC:SEV1 -46.179319   -0.4187777   13.83248124     7.3309140
0.1584383
LG10TIME:SEV1 -25.720827    0.7452832    7.33091403     9.0176263     -
0.2213634
LG10CONC:SEV2  -0.519094 -117.2769304    0.15843830    -0.2213634
33.7119235
LG10TIME:SEV2   0.099212  -62.0189016   -0.05647271     0.1892932
16.8711046
                   LG10TIME:SEV2
SEV1             0.09921199
SEV2           -62.01890157
LG10CONC:SEV1   -0.05647271
LG10TIME:SEV1    0.18929315
LG10CONC:SEV2   16.87110465
LG10TIME:SEV2   19.83371985

Analysis of Deviance Statistics:

Generalized R-squared: 0.509

         DF Deviance Mean.Dev Gen.F pvalue
Model     4 173.5442   43.386 6.729  7e-04
Residual 26 167.6454    6.448
```

```
Total     30 341.1896

################################################################
Note: About  50.9 % of the variation in the response is accounted
for by the explanatory variables in the current model fit.

The p-value of the model fit is <= 0.05. This is generally considered
significant, indicating that the current model fit is acceptable.
################################################################




*************************** ERC summary ***************************


< ERC10 at specific time point(s) >
 : One-sided 95% Lower bound and One-sided 95% Upper bound

* Risk Type : extra risk

*   ERC10 at SEV1
             ERC10 Lower Bound of ERC10 Upper Bound of ERC10
 1 Hours 1818.2862           1626.4862            2032.7038
 4 Hours  753.2313            637.7297             889.6519
 8 Hours  484.7987            368.6566             637.5304
24 Hours  241.1305            153.4754             378.8483


These One-sided 95% Lower bound and One-sided 95% Upper bound
Confidence Intervals are equivalent to
the lower bound and upper bound of Two-sided 90% Confidence Intervals

  *******************************************************************

File: chemx1.txt    Stratum: SEV1
```

If filtering is not desired, respond with <CR> to go to the next option. To filter data, the user first keys in the variable(s) from the list of variables CatReg displays on the screen. When a variable is chosen for filtering, the program responds with all the values for that variable in the entire data file. The user selects the values of the variable to filter. It may be noted that this option is also useful for scanning a particular field to see what values have been observed. Unexpected values may indicate a problem with the input data file (e.g., errors in data entry).

Clustering. An input file may consist of subsamples of data from common sources that causes them to be more similar to each other than to observations from another source. For example, suppose there are reports from three different "identical" experiments conducted at three different laboratories. The data from each laboratory may be considered a cluster because of the following four reasons. (1) There are likely some differences among laboratories in the way subjects were fed, their animal suppliers, the age of the subjects, the conditions under which

TABLE 5-3. TEXT FILE OF ERC ESTIMATES FROM EXAMPLE IN TABLE 5-1.

Time	ERC10	sd	LB95	UB95
1.25	1577.81	84.42	1444.89	1722.97
1.3	1541.64	79.75	1415.89	1678.57
1.34	1506.3	75.57	1387	1635.87
1.39	1471.77	71.87	1358.18	1594.86
1.45	1438.03	68.65	1329.43	1555.5
1.5	1405.06	65.91	1300.72	1517.78
1.56	1372.85	63.65	1272.05	1481.64
1.61	1341.38	61.84	1243.42	1447.06
1.67	1310.63	60.46	1214.86	1413.96
1.74	1280.58	59.49	1186.37	1382.28
1.8	1251.23	58.89	1158.01	1351.95
1.87	1222.54	58.63	1129.82	1322.88
1.94	1194.52	58.65	1101.84	1294.99
2.01	1167.13	58.93	1074.12	1268.2
2.08	1140.38	59.41	1046.72	1242.41
2.16	1114.23	60.07	1019.68	1217.56
2.24	1088.69	60.87	993.03	1193.56
2.32	1063.73	61.78	966.82	1170.36
2.41	1039.35	62.76	41.08	1147.88
2.5	1015.52	63.8	915.82	1126.08

```
 ** Stratum :SEV1

 ** Risk Type :extra risk

 ** Above UB95 and LB95:
    These one-sided 95% lower bound and one-sided 95% upper bound
confidence intervals are equivalent to the lower bound and upper bound of
two-sided 90% confidence intervals.
```

the subjects have been taken off the study, the protocol for histopathology (or just different histopathologists), etc. (2) The differences can be viewed as random effects by thinking of the specific laboratories as a random sample among a population of laboratories. (3) The magnitude of the differences among the specific laboratories may vary. (4) The laboratories are reasonably homogeneous (i.e., there is not one or more of them that is unreliable or consistently different in some way from the others).

As a second example, suppose that an experiment is conducted wherein pregnant female rats are exposed to a toxic substance. Each rat gives birth to a litter and the pups are examined for specific health effects. Each litter could be considered a cluster sampleClustering is necessary whenever there is reason to suspect that batches of data are correlated (i.e., when the design of the study involves cluster sampling). The cluster variable should identify each batch of correlated data uniquely. Cluster labels might be text identifiers, identification numbers, combinations of variables, etc. The only requirements are that observations from the same

cluster have the same cluster label, and those from different clusters have different cluster labels. CatReg queries for the names of variables to use in clustering the data. If no cluster variables are specified, the program treats all data as being independent.

CatReg assumes that responses from the same cluster are correlated, whereas observations from different clusters are independent. It adjusts for the cluster sampling effect using the method of generalized estimating equations (GEE). The cluster adjustment affects standard errors, confidence limits and hypothesis tests (p-values), but it does not affect parameter estimates or the deviance (a statistic used to measure the fit of exposure-response curve to the data). For technical background on GEE, see Simpson et al. (1996b) and Diggle et al. (1994). It also may be noted that cluster sampling invalidates the large sample F distribution of the generalized F-statistic. However, it is common practice to compute F as a rough guideline (see Venables and Ripley, 1994, p. 187). In any case, the R^2 statistic gives an idea of how much variation in the response is accounted for by the explanatory variables (see Section 7 for information on how F and R^2 are computed). Ignoring clusters of observations typically leads to underestimation of variability in estimates and confidence bounds that are inappropriately narrow.

Stratifying. Stratification is a way of allowing one or more of the regression parameters (intercept, coefficient for concentration ("mg/m3"), and coefficient of time ("Hours")) to change when a specified variable changes value. For example, instead of assuming a common intercept parameter for three different species, stratification of intercept on the variable Species adds two more intercept parameters so there is one for each species. Stratification of the intercept on three species defines three subgroups (strata) of data, one for each species. The same parameter can be stratified on more than one variable. For example, the intercept might be stratified on both Species and Target. If there are two target organs for each of three species, then there are six strata, each corresponding to a distinct combination of species and target. CatReg will provide six intercept estimates, one for each species-target combination. To stratify the intercept on Species and Target, enter those two variable names in response to CatReg's query on whether to stratify the intercept (enter both on the same line, separated by a space, or on separate lines). In the same way, CatReg queries for variables on which to stratify the coefficients of concentration and time.

Stratification is often conducted to test if the value of a regression parameter (e.g., the intercept) is the same for two or more values of a variable (e.g., Species). The user typically wants to produce an exposure-response curve that is suitably "accurate" by taking account of different parameter values that may occur between species, endpoints, etc., but that achieves that objective as simply as possible (i.e., with the minimal number of parameters). Stratification can be a way of implementing toxicological considerations, as the following example illustrates.

Suppose the data contain mortality results from experiments using rats and mice. The basic explanatory variables are atmospheric toxicant concentration and duration of exposure. The response score equals 0 for surviving animals and 1 for animals that died. In this case, the maximum severity score is $S = 1$. Assuming C and T enter the model logarithmically, the basic model has the form

$$L[\Pr(Y = 1 | C, T)] = \alpha_1 + \beta_{1*} \log_{10} (C) + \beta_{2*} \log_{10} (T),$$

where L is the link function (the inverse function of H in Eq. 4-1a - see Appendix B). Logarithmic scaling is typical when the explanatory variables range over two or more factors of 10. Because of different rates of respiration and metabolism among rats and mice, it may be reasonable to assume the internal dose for rats should be rescaled compared to that of mice. One possibility is to assume that a concentration of C for rats is equivalent to a concentration of kC for mice, where k is common to all mice in the study. Then, for the mice,

$$L[\Pr(Y = 1 | C, T)] = \alpha_1 + \beta_{1*} \log_{10} (kC) + \beta_{2*} \log_{10} (T).$$
$$= [\alpha_1 + \beta_{1*} \log_{10} (k)] + \beta_{1*} \log_{10} (C) + \beta_{2*} \log_{10} (T).$$

This shows that the mice, in effect, have a different intercept than do the rats, namely $\alpha_1^{MU} = \alpha_1^{RT} + \beta_{1*} \log_{10}(k)$, where MU and RT refer to mouse and rat parameters, respectively. By stratifying the intercept parameter, the data are allowed to determine the estimate of the conversion factor k. Whether k is significant would be determined by testing $\alpha_1^{MU} = \alpha_1^{RT}$.

Scaling. If the logarithmic option is chosen for concentration (mg/m3) and there is no stratification, the user is asked Assume zero background risk (i.e., response cannot occur at zero concentrations)?(y). If the default (y) is entered, then the implied probability of an adverse response at zero concentration is zero and observations at zero concentration are uninformative (treated the same as if they were filtered out). If the response is "n" then CatReg adds a hypothetical background concentration to the administered

concentrations, denoted as the parameter γ (gamma) that is estimated by maximum likelihood and displayed in the summary table of parameter estimates. For example, an experimental concentration of 50 mg/m3 is treated as an observation at concentration $(50 + γ)$ mg/m3, where γ is estimated from the data simultaneously with the other model parameters (see Section 5.4). If the logarithmic option is chosen for time (Hours), the implied probability of an adverse response at zero time is zero and observations at zero time are uninformative (treated the same as if they were filtered out).

Linking. A link function is a function applied to the exposure-response curve to transform it to a simple linear relationship in concentration and duration. By also transforming the observed responses, the link function reduces the mathematical complexity of estimating the parameters. The parameter estimates then are substituted into the (untransformed) exposure-response curve. There is a unique link function for each of the three choices of exposure-response curves provided by CatReg (i.e., to the three different formulations for the function H in Eq. 4-1a,b). CatReg simply asks the user to select one of those link functions: logit, probit, or cloglog (equivalent to selecting the logistic, normal, or Gumbel probability function for H, respectively). The link function may be derived from a basic assumption that the ordinal severity score corresponds to exceeding an underlying toxic response threshold. See Appendix B for an example and further discussion.

A comparison of how well the different link functions fit the data may be assessed using the AIC (Akaike Information Criteria, Akaike (1974)). A link with a smaller AIC provides a closer fit to the data. The AIC for different link functions may be compared only if there are no changes in the data or in use of the data filtering and stratification options. Otherwise the differences in AIC may result from the changes.

Model. There are two choices of models, Model 1 (cumulative odds model) and Model 2 (unrestricted cumulative model), as described in Section 4.1. A conditional odds model and unrestricted conditional model are not currently provided. Equation (4-1b) describes the unrestricted cumulative model. The cumulative odds model is the same except that the parameters $β_{1s}$ and $β_{2s}$, the coefficients of concentration (or log concentration) and time (or log time), respectively, do not depend on s, the severity level, i.e., Model 1 is a simplification of Model 2 in which only the intercept term can vary across severity levels, not the coefficients of concentration or time (a restriction called *parallelism*). In other words, the cumulative odds

37

model (Model 1) states that the probability that a severity level "s" or higher will occur at a given concentration (C) and time (T) is given by the exposure-response curve (logistic, normal, or Gumbel, determined by the choice of link function), and that the intercept parameters may differ by severity level (i.e., a different intercept for each severity level), but the coefficients for C and T do *not* differ by severity.

The restriction of parallelism implies that the exposure- response curves for different severity levels will not cross. Crossing would indicate a contradiction (e.g., that the probability of severity 1 or greater is less than the probability of severity 2 or greater) for some values of C and T. The unrestricted cumulative model (Model 2) differs by estimating separate coefficients for C and T at each severity level. Using this option, it is possible that the severity curves might cross, but often the values for C and T at which there would be a contradiction are well outside the range of values of interest, so the model can be used to make empirical risk estimates.

The user has an option to fit Model 2 and then test the hypothesis of parallelism by using the function `parallel.test()` (see Section 6.1) or fitting both Models 1 and 2 and then selecting the outcome with the lower AIC value. Model 2 requires more data than Model 1 to converge on a solution because more parameters are estimated, and Model 2 will have a lower deviance because the larger number of parameters will produce a better fit to the data. Model 2, however, requires more parameters. The AIC takes both into consideration. Having more parameters in the model decreases the first part of the AIC (-2 times the log of the minimized value of likelihood function) but at the cost of increasing the second part of the AIC (+ 2 times the number of parameters in the model).

Worst Case. CatReg provides an option to do a worst-case analysis that only appears on the screen when there is at least one record in the input file that contains censored data: `Worst-case analysis of interval-censored responses? (n)`. In a worst case analysis, censored responses are treated as occurring at their highest possible (worst) severities. Although the graphical presentation of censored points will not change in a worst-case analysis, higher estimates of risk will be produced than those for the corresponding censored analysis. Comparison of risk estimates from the two methods provides an indication of the sensitivity of the results to the severity scoring. When the worst case option is selected, the output file contains the line: `Type of analysis: Worst-case`. When the worst-case option is not selected, the output line is: `Type of analysis: Censored`.

5.3. Setup options—Examples

As discussed in Section 5.1, Table 5-1 shows the screen display for an example using input file chemx.csv, with data for Species = MU and Target = L filtered (i.e., excluded), leaving only data from the single experiment where Species = RT and Target = C. The output file is shown in Table 5-2 and the ERC estimates that were written to a text file are in Table 5-3. Model 2 was fit to the data, so there are separate intercept parameters for severity levels 1 and 2, the same as if Model 1 were used to fit the data, but also separate concentration coefficients and time coefficients for each severity level. In the table of parameter estimates, `SEV1` and `SEV2` correspond to the parameters α_1 and α_2 of Eq. 4-1b, `LG10CONC:SEV1` and `LG10CONC:SEV2` correspond to β_{11} and β_{12}, and `LG10TIME:SEV1` and `LG10TIME:SEV2` correspond to β_{21} and β_{22}. The logit model was used. The probit link would decrease the AIC slightly, from 179.6 to 178.6, and the cloglog link would increase it to 180.5. For this example, there is only a small difference in the three model choices using the AIC as a statistical measure for comparison.

For illustration, we continued with the logit link and modified the setup options to select Model 1, the cumulative odds model. Only the intercept parameter was stratified on severity, leaving just four parameters in the model, `SEV1, SEV2, LG10CONC,` and `LG10TIME`, corresponding to parameters α_1, α_2, β_1, β_2, respectively. The AIC dropped to 177.3, favoring Model 1, the simpler model, over Model 2, although the difference is not very large.

We then considered input file chemx.csv without filtering the data. There are data for two species (RT and MU) and two target organs (L and C), and the intercept, concentration, and duration, all three, were stratified on both species and target. Selecting Model 2 produced an error message,

```
 Error: Sev1 < Sev2! Estimates of severity parameters do not
satisfy constraint on order of parameters. Incorrectly ordered
severity estimates is evidence of too many severity levels in
the data. This run will terminate.
```

The error message indicates that the exposure-response curves for the two severity levels cross, at least for the reference stratum which is MU:C. That suggests combining severity levels 1 and 2, but it is not clear whether that should apply to all target-species combinations. To address that question, CatReg was run separately for the subsets of data for each target-species

combinations (MU:C, MU:L, RT:C, RT;L) and the same message occurred in all but the data for RT:C. For RT:C, a test of the hypothesis SEV1 = SEV2 was not significant (p = 0.24, using the CatReg function `partest()`, discussed in Section 6.1). Thus, it would be reasonable in this example, from the statistical perspective, to join severity level 2 with severity level 1, eliminating severity level 2 (see discussion of `join()` in Section 10.4). For illustration, however, we retained severity level 2 and repeated the same CatReg run except with Model 1 (cumulative odds model). Model 1 reduces the number of parameters to be estimated but at the expense of the assumption of parallelism (i.e., that the coefficients of concentration and duration do not vary by severity level). Fitting the data to Model 1 produced an AIC value of 756.0 for a model with 13 parameters. The output file, displayed as Table 5-4, will be discussed.

In terms of Equation 4-1a, Model 1 with two severity levels but *without stratification* would have the following parameters: intercept terms (α_1, α_2 for severity levels 1 and 2), a concentration coefficient (β_1), and a time coefficient (β_2). *With stratification* of the intercept, concentration coefficient, and time coefficient, all on both species and target, at severity level 1 there would be four estimates for each parameter, one for each species/target combination (species MU and RT with targets L and C) as in Table 5-4. The parameter estimates in Table 5-4 are interpreted as follows, for severity level 1. The stratum MU:C (mice/central nervous system) is the reference group. No intercept parameter is estimated for this group. Its intercepts are given by SEV1 for severity level 1 and SEV2 for severity level 2. The MU:C:INTERCEPT row is merely to identify the reference group for the intercept parameter. "NA" under the headings for Z-test and p-value means "not applicable" or "not available." Other intercept parameters are increments relative to the reference parameter.

The intercept term for MU:C, at severity level 1 (SEV1), can be written as SEV1 + MU:C:INTERCEPT, estimated as -61.56 + 0. Similarly, the intercept term for MU:L is SEV1 + MU:L:INTERCEPT, estimated as 61.56 - 2.34, etc. The intercept terms and their estimates at severity level 2 are obtained the same way except with SEV2 replacing SEV1. The Z-test for MU:L:INTERCEPT is a test of the null hypothesis that MU:L:INTERCEPT is zero. Since

TABLE 5-4. OUTPUT FILE FROM CATREG. INPUT FILE: CHEMX.CSV. MODEL 1. LOGIT. SCALES: LOG10. STRATIFICATION: INTERCEPT, CONCENTRATION, DURATION, ALL ON SPECIES AND TARGET. ASSUMED ZERO BACKGROUND RISK.

```
CatReg - R Beta Version 1.1 (12/07/05)
```

```
Source data file: chemx.csv

Type of analysis: Censored

Input file    : chemx.csv
Filtered data: none
Model         : cumulative odds model
Link          : logit
Clustering    : none
Message       :
Iterations    : 49 28
Deviance      : 729.9682
Residual DF   : 115
AIC           : 755.9682

Scale:
     Concentration: log10( mg/m3 )
     Duration     : log10( Hours )

Stratification:
     Intercept     : Species Target
     Concentration: Species Target
     Duration      : Species Target

Coefficients:
                   Estimate Std. Error   Z-Test=0     p-value
SEV1             -61.5577058  6.729186  -9.1478681    0.00001
SEV2             -65.7115174  6.837497  -9.6104632    0.00001
MU:C:INTERCEPT     0.0000000  0.000000         NA         NA
MU:L:INTERCEPT    -2.3399958  9.126294  -0.2564015    0.79764
RT:C:INTERCEPT   -18.5307081  9.278765  -1.9971093    0.04581
RT:L:INTERCEPT    -0.6519446  8.939393  -0.0729294    0.94186
MU:C:LG10CONC     18.2050464  2.017313   9.0244020    0.00001
MU:L:LG10CONC     19.2629512  2.182794   8.8249068    0.00001
RT:C:LG10CONC     23.7779909  2.293245  10.3687104    0.00001
RT:L:LG10CONC     18.5222850  2.099372   8.8227750    0.00001
MU:C:LG10TIME     17.9322449  2.018420   8.8842976    0.00001
MU:L:LG10TIME     19.8904023  2.183654   9.1087686    0.00001
RT:C:LG10TIME     16.0546314  2.062325   7.7847228    0.00001
RT:L:LG10TIME     19.5678918  2.120738   9.2269266    0.00001

Analysis of Deviance Statistics:

Generalized R-squared: 0.472

          DF  Deviance Mean.Dev Gen.F pvalue
Model     11  653.6035   59.418 9.361      0
Residual 115  729.9682    6.348
Total    126 1383.5717

############################################################################
Note: About  47.2 % of the variation in the response is accounted
for by the explanatory variables in the current model fit.
The p-value of the model fit is <= 0.05. This is generally considered
significant, indicating that the current model fit is acceptable.
############################################################################


**************************** ERC summary ****************************


< ERC10 at specific time point(s) >
 : One-sided 95% Lower bound and One-sided 95% Upper bound
```

```
* Risk Type : extra risk

*  ERC10 at SEV1:MU:C
            ERC10 Lower Bound of ERC10 Upper Bound of ERC10
 1 Hours 1822.4717          1642.4437          2022.2325
 4 Hours  465.1817           370.6169           583.8753
 8 Hours  235.0193           164.4600           335.8511
24 Hours   79.6401            45.1044           140.6193


*  ERC10 at SEV1:MU:L
            ERC10 Lower Bound of ERC10 Upper Bound of ERC10
 1 Hours 1596.1165          1440.8915          1768.0637
 4 Hours  381.4115           299.0930           486.3861
 8 Hours  186.4482           128.8548           269.7839
24 Hours   59.9647            33.7120           106.6613


*  ERC10 at SEV1:RT:C
            ERC10 Lower Bound of ERC10 Upper Bound of ERC10
 1 Hours 1886.9845          1730.8095          2057.2515
 4 Hours  740.0554           646.4460           847.2199
 8 Hours  463.4599           371.5325           578.1326
24 Hours  220.7320           153.5434           317.3214


*  ERC10 at SEV1:RT:L
            ERC10 Lower Bound of ERC10 Upper Bound of ERC10
 1 Hours 1737.8489          1567.8826          1926.2405
 4 Hours  401.7584           313.4229           514.9905
 8 Hours  193.1708           131.6381           283.4662
24 Hours   60.5182            33.0944           110.6668


These One-sided 95% Lower bound and One-sided 95% Upper bound
Confidence Intervals are equivalent to
the lower bound and upper bound of Two-sided 90% Confidence Intervals

    *****************************************************************
```

MU:L:INTERCEPT is the incremental difference from SEV1, which is the intercept for MU:C, the hypothesis tested by the Z-test is that the intercepts for liver and central nervous system are equal in mice. Similarly, the Z-test for RT:C:INTERCEPT is a test of the null hypothesis that the intercepts for rats and mice is the same for the central nervous system. Other hypotheses can be tested using the command partest(), discussed in Section 6.1.

For both severity levels 1 and 2, the estimates of the concentration coefficients and the time coefficients for each combination of species and target can be read directly from the table. For example, MU:C:LG10CONC refers to the coefficient of concentration (β_1) for mice and the central nervous system; MU:C:LG10TIME refers to the coefficient of time (β_2) for mice and the central nervous system, for both severity levels 1 and 2, etc. The Z-tests for coefficients of

concentration and time are tests of the null hypothesis that the parameter is zero, e.g., the Z-test for `MU:L:LG10CONC` tests the hypothesis that the coefficient of concentration for a liver effect in mice is zero, i.e., that the substance tested has no adverse effect in the range of concentrations tested. In general, the CatReg convention for stratification of parameters is:

1. The stratification of the intercept parameter assigns a reference category, so that other intercept parameters are increments relative to this category.
2. The stratification of the concentration and time parameters results in separate parameters for each stratum.

Choosing linear scales for concentration and time instead of logarithmic scales increased the AIC for the example to 764, so we continued to use the log scales. When only the intercept was stratified on species and target, the AIC remained about the same (755) so we continued with the simpler model, i.e., stratifying only the intercept on species and target. The output file is displayed in Table 5-5. We might now ask if stratification on the intercept can be reduced further. Stratification of intercept on species or target alone, however, caused a large increase in AIC (to 817.5 and 769.2, respectively), so no further simplification was achieved.

5.4. Nonzero Background Response

As noted previously when C (concentration) is log-transformed, the probability of an adverse effect of level 1 or higher, expressed in Eq. 4-1a,b, approaches zero as concentration approaches zero. If there is a positive probability of an adverse effect even when concentration is zero, i.e., so-called background response not attributable to exposure, then the user can modify the probability function in Eq. 4-1a,b. When the user chooses the log scale for concentration, an option appears on the screen: `Assume zero background risk (i.e., response cannot occur at zero concentration)?(y)`. The exposure-response curve in

TABLE 5-5. OUTPUT FILE FROM CATREG. INPUT FILE:CHEMX.CSV. MODEL 1. LOGIT. SCALES: LOG10. STRATIFICATION: INTERCEPT ON SPECIES AND TARGET. ASSUMED ZERO BACKGROUND RISK.

```
CatReg - R Beta Version 1.1 (12/07/05)

Source data file: chemx.csv

Type of analysis: Censored

Input file    : chemx.csv
Filtered data: none
Model         : cumulative odds model
Link          : logit
Clustering    : none
Message       :
Iterations    : 41 14
Deviance      : 741.0266
Residual DF   : 121
AIC           : 755.0266

Scale:
      Concentration: log10( mg/m3 )
      Duration     : log10( Hours )

Stratification:
      Intercept    : Species Target
      Concentration:
      Duration     :

Coefficients:
                    Estimate Std. Error    Z-Test=0     p-value
SEV1            -66.2375249  4.2947663  -15.422847     0.00001
SEV2            -70.3128027  4.4629521  -15.754774     0.00001
MU:C:INTERCEPT    0.0000000  0.0000000          NA          NA
MU:L:INTERCEPT    1.5443003  0.2803179    5.509103     0.00001
RT:C:INTERCEPT   -0.8580588  0.2701988   -3.175658     0.00149
RT:L:INTERCEPT    0.7735915  0.2703681    2.861253     0.00422
LG10CONC         19.6220725  1.2858838   15.259600     0.00001
LG10TIME         18.0738765  1.2265874   14.735091     0.00001

Variance:
                        SEV1         SEV2 MU:L:INTERCEPT RT:C:INTERCEPT
SEV1             18.4450174  19.1496806    -0.32566226     0.13047885
SEV2             19.1496806  19.9179410    -0.34292950     0.14029845
MU:L:INTERCEPT   -0.3256623  -0.3429295     0.07857815     0.03192034
RT:C:INTERCEPT    0.1304789   0.1402984     0.03192034     0.07300738
RT:L:INTERCEPT   -0.1730734  -0.1814300     0.03822081     0.03365877
LG10CONC         -5.5095904  -5.7215951     0.08581225    -0.04903025
LG10TIME         -3.4384286  -3.6292401     0.08062509    -0.04554744
                RT:L:INTERCEPT      LG10CONC       LG10TIME
SEV1              -0.17307335   -5.50959042    -3.43842860
SEV2              -0.18143003   -5.72159505    -3.62924010
MU:L:INTERCEPT     0.03822081    0.08581225     0.08062509
RT:C:INTERCEPT     0.03365877   -0.04903025    -0.04554744
RT:L:INTERCEPT     0.07309891    0.04053521     0.04091319
LG10CONC           0.04053521    1.65349721     0.97163480
LG10TIME           0.04091319    0.97163480     1.50451662

Analysis of Deviance Statistics:

Generalized R-squared: 0.464

        DF  Deviance Mean.Dev  Gen.F pvalue
```

44

```
Model       5  642.5451  128.509 20.984        0
Residual 121  741.0266    6.124
Total     126 1383.5717
```

```
##############################################################################
Note: About  46.4 % of the variation in the response is accounted
for by the explanatory variables in the current model fit.

The p-value of the model fit is <= 0.05. This is generally considered
significant, indicating that the current model fit is acceptable.
##############################################################################


***************************** ERC summary *****************************


< ERC10 at specific time point(s) >
 : One-sided 95% Lower bound and One-sided 95% Upper bound

* Risk Type : extra risk

*   ERC10 at SEV1:MU:C
             ERC10 Lower Bound of ERC10 Upper Bound of ERC10
 1 Hours 1835.2147           1729.5769           1947.3045
 4 Hours   511.8349            460.2481            569.2037
 8 Hours   270.3033            229.7032            318.0794
24 Hours    98.2597             75.9918            127.0528


*   ERC10 at SEV1:MU:L
             ERC10 Lower Bound of ERC10 Upper Bound of ERC10
 1 Hours 1531.0334           1442.0397           1625.5193
 4 Hours   426.9998            380.3650            479.3523
 8 Hours   225.5014            189.8653            267.8260
24 Hours    81.9734             62.8295            106.9505


*   ERC10 at SEV1:RT:C
             ERC10 Lower Bound of ERC10 Upper Bound of ERC10
 1 Hours 2029.6265           1911.0637           2155.5451
 4 Hours   566.0556            511.4375            626.5066
 8 Hours   298.9376            255.2595            350.0897
24 Hours   108.6687             84.4407            139.8483


*   ERC10 at SEV1:RT:L
             ERC10 Lower Bound of ERC10 Upper Bound of ERC10
 1 Hours 1675.9548           1579.5678           1778.2236
 4 Hours   467.4179            418.1326            522.5124
 8 Hours   246.8464            208.6956            291.9714
24 Hours    89.7327             69.0508            116.6092



These One-sided 95% lower bound and one-sided 95% upper bound
confidence intervals are equivalent to the lower bound and upper bound of
two-sided 90% confidence intervals

    ****************************************************************************
```

Eq. 4-1a,b is modified by adding a hypothetical background concentration level, γ (gamma), to the administered concentration C given in the input file (as variable mg/m3). The parameter γ is estimated by maximum likelihood simultaneously with the other parameters and the result is added to the summary table of parameter estimates. If the user responds with "y" to assume zero background risk, then the data records where concentration is zero are non-informative and CatReg ignores them (i.e., effectively filters those data). That reduces the total degrees of freedom compared to a response of "n" which uses the data where concentration is zero.

Table 5-6 lists an input file that was generated by simulation using log- transformed C and hypothetical background concentration $\gamma = 20$ mg/m3. Evidence of non-zero background risk is apparent in the occurrence of both severity levels 1 and 2 when mg/m3 = 0. The other parameter values used for simulation are $\alpha_1 = -7.5$, $\alpha_2 = -9.5$, $\beta_{11} = 2.5$, $\beta_{12} = 2.0$, $\beta_{21} = 2.0$, $\beta_{22} = 1.8$. The data were simulated for the unrestricted cumulative model (Model 2) with the logit link, and log-transformed T (exposure duration). Responding "n" to the option to assume zero background risk during execution informs CatReg to add the parameter γ. The output file is displayed in Table 5-7. The evidence of background risk is not significant in this data set. The estimate of gamma is 16.0 with standard error 28.1, which is not significantly different from zero.

This example suggests that one might need a substantial background effect for it to be significant, at least for small sample sizes. The current example consists of treatment groups of size 10 each at 16 exposures (concentrations of 0, 80, 180, and 480; durations of 20, 100, 200, and 500 hours). To examine the effect of sample size further, the same example was repeated but with treatment group sizes of 50, 100, and 5,000. The estimate, standard error, and significance level of gamma, for sample sizes of 50, 100, and 5,000, respectively, were: (31.9, 20.7, 0.12), (12.6, 7.7, 0.10) and (21.5, 1.6, and $<10^{-5}$). For this example, estimates of gamma appear to become statistically significant and to converge to the neighborhood of the correct value, 20, very slowly. When 1000 datasets were simulated with treatment groups of size 10, the median estimate of gamma was 19.7 (20.0 was the value of gamma used for simulation), but the standard error of the estimates was high (69.3).

TABLE 5-6. INPUT FILE BACKGD10.CSV.

mg/m3	Hours	SevLo	Nsub	Incid
0	20	0	10	8
0	20	1	10	2
0	20	2	10	0
0	100	0	10	6
0	100	1	10	4
0	100	2	10	0
0	200	0	10	5
0	200	1	10	4
0	200	2	10	1
0	500	0	10	3
0	500	1	10	7
0	500	2	10	0
80	20	0	10	4
80	20	1	10	6
80	20	2	10	0
80	100	0	10	2
80	100	1	10	8
80	100	2	10	0
80	200	0	10	1
80	200	1	10	8
80	200	2	10	1
80	500	0	10	0
80	500	1	10	4
80	500	2	10	6
180	20	0	10	3
180	20	1	10	6
180	20	2	10	1
180	100	0	10	3
180	100	1	10	7
180	100	2	10	0
180	200	0	10	0
180	200	1	10	5
180	200	2	10	5
180	500	0	10	1
180	500	1	10	4
180	500	2	10	5
480	20	0	10	0
480	20	1	10	8
480	20	2	10	2
480	100	0	10	1
480	100	1	10	8
480	100	2	10	1
480	200	0	10	0
480	200	1	10	2
480	200	2	10	8
480	500	0	10	0
480	500	1	10	2
480	500	2	10	8

TABLE 5-7. CATREG SUMMARY FROM OUTPUT FILE OF CATREG. INPUT FILE: BACKGD10.CSV. MODEL 2. LOGIT. SCALES: LOG10. ESTIMATED BACKGROUND RISK.

```
Input file   : backgd10.csv
Filtered data: none
Model        : unrestricted cumulative model
Link         : logit
Clustering   : none
Message      :
Iterations   : 26 10
Deviance     : 239.8087
Residual DF  : 25
AIC          : 253.8087

Scale:
     Concentration: log10( mg/m3 )
     Duration     : log10( Hours )

Stratification:
     No Stratification on Intercept, Concentration and Duration.

Coefficients:
               Estimate Std. Error   Z-Test=0   p-value
SEV1           -6.279480  3.5472283 -1.7702497  0.07669
SEV2          -13.394898 33.7617732 -0.3967475  0.69155
LG10CONC:SEV1   2.305462  1.4596636  1.5794473  0.11423
LG10TIME:SEV1   1.567190  0.4607761  3.4011975  0.00067
LG10CONC:SEV2   2.747634  0.6149344  4.4681741  0.00001
LG10TIME:SEV2   2.777898  0.6190206  4.4875700  0.00001
Gamma          15.956845 28.1522306  0.5668057  0.57085
```

The smallest positive value of concentration in an input file is used as a practical boundary on gamma. If the estimate of gamma is set to the boundary value, a message appears in the output: Warning: Gamma hit its maximum bound!!!. The estimate shown for gamma, and the other parameters in the summary table of estimates, are not maximum likelihood in that case, and the user is advised to consider the setup option that assumes the background risk is zero.

5.5.   Data Requirements and Error Messages

Certain minimal data requirements need to be satisfied in order to estimate the categorical regression model. There needs to be at least one response in each severity category. If some categories are completely absent, then categories will need to be combined (see discussion of the "join" function, Section 10.4). If both concentration and duration effects are to be modeled, then both C and T need to be varying in the data. If T is constant, then a reduced model is fit so that T is dropped from the model. There are additional technical limitations on the complexity of the

48

model relative to the data. The most obvious is that the model cannot include more parameters than the number of independent observations. Less obvious problems sometimes occur if certain variables are redundant. These problems are revealed by R error messages indicating that there are redundant variables in the model. This means that at least one variable in the model can be expressed in terms of the others. Reducing the number of stratification variables usually will solve the problem. Sometimes, R will return a "failure to converge" message. This is an indication that there are too many variables in the model, and that the model needs to be simplified. This problem relates to the number of variables needed to completely isolate the different severity categories. The solution is to remove one or more variables from the model.

Unless the user is familiar with the R programming language and its use in the CatReg manuscript files, the error messages offered by CatReg are rather indecipherable. Although it is beyond the scope of this manual to educate the user in the R programming language and syntax, some common error messages and proposed solutions will be offered. The format for this discussion will follow the order of CatReg inquiries shown in Table 5-1 and address the error messages that may occur as the analysis progresses. In this section, the error messages are printed in `this font (Courier New)` for easy recognition.

After the initial `catreg()` command, `Error in file(file, "r") : unable to open connection. In addition: Warning message: cannot open file `catreg.R'` indicates that the CatReg script files have not been installed properly. Check to make sure that the script files have been installed in the directory designated in what R refers to as the "Start in" directory. You can type in `getwd()` to verify this.

After the user enters the name of the input file, `Error in count.fields (infile, sep = ",") : cannot open the connection. In addition: Warning message: cannot open file `testcase.csv' for reading` means that the data file cannot be found. Check to make sure the name of the data file has been accurately typed and that the data file is contained in the "Start in" directory. If an error message such as `Error in askfordata(infile): >>>> Records 2, 6, 10, 12, 15, 18, 19, 21, 22, 24, 38, 39, 51, 53, 55, 61, 63, 67 have missing or extra fields!` appears, the format of the data file is wrong. Check to make sure the appropriate column headings are used (see Section 3.2) and that all data, including column headings, are separated by commas. If a required field of data is missing, an error message is

produced to that effect, e.g., `Error in askfordata(infile) : Required input variable SevLo was not found.`

If the value of "Incid" for the records in a treatment group (also called a "data entry") do not sum to the common value of "Nsub" in those records, then an error message occurs giving the approximate record number of the error, e.g.

`Error detected in input, on or near line 4.`

`Total incidence for a data entry does not equal number of subjects. CatReg will close (disregard the following line).` The data input file needs to be corrected.

In proceeding with CatReg's inquiries for the various setup options, a mistyped response will invoke `Enter an item from the menu, or 0 to exit.` Simply type the correct response after CatReg repeats the query.

If an error is made in typing the response to `Worst-case analysis of interval-censored responses (n)?` CatReg proceeds under the assumption that the default response, "n" (for no), has been entered. Thus, it accepts "y" as yes and anything else as "no".

Sometimes "NaN" (not a number) is given for some of the model deviance iterations. It is printed when CatReg is searching for the solutions to the parameter estimates in early iterations of the program. CatReg tries to compute the deviances for those solutions and finds they are undefined (NaN). Once a solution is determined for runs like this, the following warning message may appear: `NAs produced in: log(likes)`(Na means "not available").

As described in the preceding section, the message `Warning: Gamma hit its maximum bound!!!` may occur when the parameter γ (gamma) is being estimated. The smallest positive value of concentration in an input file is used as a practical boundary on gamma. The estimate shown for gamma, and the other parameters in the summary table of estimates, are not maximum likelihood in that case, and the user is advised to consider the setup option that assumes the background risk is zero.

An error message occurs if a coefficient of concentration or time is negative, e.g., `Error: TIME is negative! Estimates of coefficient parameters do not satisfy non-negativity constraint on the parameters. A`

negative estimate is evidence of no effect. This run will
terminate. The user needs to modify the run. The data may be indicating there is no effect or there may just be too many parameters in the model.

Similarly, the estimates of severity intercepts may violate the order constraint, resulting in a message such as Sev 1 < Sev 2! Estimates of severity parameters do not satisfy constraint on order of parameters. Incorrectly ordered severity estimates is evidence of too many severity levels in the data. This run will terminate. The user needs to modify the run, such as combining severity levels with the join function (see Section 10.4).

Most error messages at this point in a CatReg run result from an attempt to calculate too many parameters with too little data. This may generate error and /or warning messages from the optimization routine (used to find maximum likelihood estimates), such as: Error in optim(transcoefs, fn = hfdeviance, gr = hfdeviance.grad, method = "BFGS",: initial value in vmmin is not finite
In addition: Warning messages:
1: Algorithm did not converge in: (if (is.empty.model(mt)) glm.fit.null else glm.fit)(x = X, y = Y,
2: fitted probabilities numerically 0 or 1 occurred in: (if (is.empty.model(mt)) glm.fit.null else glm.fit)(x = X, y = Y,
3: NaNs produced in: log(x),
or an error message that the covariance matrix of estimates is singular: Error in solve.default(Hessian):system is computationally singular: reciprocal condition number = …

In general, when errors occur during the calculation of parameter estimates, the user should change the link function if the cloglog link function was used or simplify the model and run it again. The model can be simplified by reducing or omitting stratification of parameters. If Model 2 (unrestricted cumulative model) is in use, the user can consider switching to Model 1 (cumulative odds model). Additionally, one can try different options for the scale of the explanatory variables concentration and duration. Dumped means that the CatReg analysis has been terminated because a solution could not be computed.

If an error is made in typing the response to `Write ERC data to text files (n)?`, CatReg proceeds under the assumption that the default response, "n", has been entered. Thus, it accepts "y" as yes and anything else as "no".

6.   TESTING HYPOTHESES

6.1.   Tests on Parameters

As discussed briefly in Section 5.1 the table of coefficient estimates shown on the screen and in the output file, illustrated in Tables 5-1, 5-2, 5-4 and 5-5, includes a simple test of the hypothesis that a parameter is zero, for each parameter in the table. Dividing the estimated coefficient by the standard error provides a Z statistic for the corresponding parameter. This statistic provides a one-degree-of-freedom test of the null hypothesis that the parameter equals zero. Under the null hypothesis, the Z statistic has an approximate standard normal distribution. The larger the sample, the better the approximation is. The p-value in the table gives the significance level of the test. A parameter that is not significantly different from zero may be considered a candidate for removal and simplification of the exposure-response curve. The p-values apply to individual parameters considered singly, however, and further testing using the function `partest` is needed to test the joint hypothesis that more than one parameter is zero, that two or more parameters are equal, or a combination of the two.

In response to the instruction `partest()`, CatReg lists all the parameters and requests the user to enter the parameters to test for removal, then lists the parameters again and asks the user to enter a group of parameters to test for equality (the default is none in both cases). The idea is to express the hypothesis to be tested as a set of constraints on the model coefficients. A test is then conducted of the hypothesis, or joint hypotheses if more than one was entered. The test is a (generalized Wald-type) chi-square test of the null hypothesis that all of the specified constraints hold. The distribution of the test statistic is derived from the sampling distribution of the estimated model coefficients, and it takes into account any cluster sampling, as described in Section 5.2.

As an example, one might like to test for a gender difference, or whether there are interspecies differences in the exposure-response. For such a test, the null hypothesis is that the specified parameters are equal. A p-value less than 0.05 is usually taken as evidence that the hypothesis should be rejected (i.e., the specified parameters are *not* equal). The parameters to be

tested as equal must, of course, be included in the exposure-response curve; this is accomplished by stratifying. For example, to test that there is no difference between species in the coefficients of concentration, the user would stratify the concentration parameter on species. When `partest` queries which coefficients to test for equality, the parameters to be tested (if any) are entered from the list of parameters provided by CatReg. To test for no difference between species in the coefficients of concentration, the user would enter the coefficient for each species. Some care needs to be used when tests involve the intercept parameters because of the way they are represented as increments relative to a reference (discussed below).

As an example, consider the previous example where Table 5-1 displays what appears on the monitor and Table 5-2 contains the output file. The data are filtered so there is only one species (RT) and one target (C) and Model 2 (unrestricted cumulative odds) is fit to the data. In Section 5.3 the same data were fit to Model 1 (cumulative odds), which differs by having single parameters for the coefficient of concentration and the coefficient of time instead of separate parameters for each severity level (referred to previously as "parallelism"). The AIC value indicated that the simpler model (Model 1) was adequate.

Another approach would have been to use `partest` to include a test of the joint hypothesis that $\beta_{11} = \beta_{12}$ and $\beta_{21} = \beta_{22}$ (in terms of the parameters of Eq. 4-1b) using `partest`. In terms of the coefficient names in Tables 5-1 and 5-2, the hypothesis is that LG10CONC:SEV1 = LG10CONC:SEV2 and LG10TIME:SEV1 = LG10TIME:SEV2. Table 5-8 shows how the example of Table 5-1 would be continued to test the hypothesis using `partest` (the user-input is in boldface). A p-value less than 0.05 (or another user-specified level of significance) is evidence that the data do not conform to the assumption of parallelism. This particular application of `partest`, i.e., to test whether Model 2 can be reduced to Model 1, has its own CatReg function. `Partest` was used for illustration but in practice the result can be obtained more simply by just entering the instruction `parallel.test()` after fitting the data to Model 2 (unrestricted cumulative model).

In conducting tests, it is necessary to be aware of which estimates are incremental changes from others, as in the case of some intercept parameters, and which are "stand alone", as this can affect how a test is constructed. For illustration, consider the output file in Table 5-5. The intercept is stratified on both species and target. MU:C:INTERCEPT is the reference, i.e., SEV1 and SEV2 apply to the central nervous system of mice. The remaining intercept terms are

TABLE 5-8. MONITOR DISPLAY OF ADDING COMMAND `PARTEST()` AT THE END OF EXAMPLE IN TABLE 5-1.

```
> partest()

 Parameters:
 SEV1 SEV2 LG10CONC:SEV1 LG10TIME:SEV1 LG10CONC:SEV2 LG10TIME:SEV2

 Enter parameters to test for removal (none):
1:
Read 0 items

 Parameters:
 SEV1 SEV2 LG10CONC:SEV1 LG10TIME:SEV1 LG10CONC:SEV2 LG10TIME:SEV2

 Enter group of parameters to test for equality (none):
1: LG10CONC:SEV1
2: LG10CONC:SEV2
3:
Read 2 items

 Parameters:
 SEV1 SEV2 LG10TIME:SEV1 LG10TIME:SEV2

 Enter another group to test for equality (none):
1: LG10TIME:SEV1
2: LG10TIME:SEV2
3:
Read 2 items

 Parameters:
 SEV1 SEV2

 Enter another group to test for equality (none):
1:
Read 0 items

Hypothesis Test Results
-----------------------

Coefficients tested for removal:   None

Coefficients tested for equality:

    Group 1 :    LG10CONC:SEV1 LG10CONC:SEV2

    Group 2 :    LG10TIME:SEV1 LG10TIME:SEV2

Test statistics:

   Chisquare      df      p-value
   1.555015        2      0.45955

######################################################################
The P Value of the equality test is > 0.05. This is generally considered not
significant, indicating that the tested parameters are equal to each other.
There is no need to keep all of them in the model.
######################################################################
```

54

incremental differences from the reference. To test whether there is no target effect on the intercept term, in executing `partest`, MU:L:INTERCEPT would be a parameter for removal and RT:C:INTERCEPT and RT:L:INTERCEPT would be a group to test for equality (to be tested jointly, in the same call to `partest`). If MU:C:INTERCEPT and MU:L:INTERCEPT were entered as a group to test for equality, a message would be returned that there is no match for MU:C:INTERCEPT.

7. ASSESSING MODEL FIT

This section focuses on the fit of the exposure-response curve rather than on its parameters. In linear regression analysis, it is common practice to consider the proportion of variation accounted for by the model, the so-called R^2 statistic, as a measure of the model's explanatory power. This statistic, which ranges in value between 0 and 1, is the ratio of the model and total sums of squares. These sums of squares, along with degrees of freedom, F-tests, and so on usually are reported in the form of an analysis of variance table. Standard texts such as Weisberg (1985) describe the use and interpretation of these statistics.

CatReg provides generalized analysis of variance and R^2 statistics for assessing the explanatory capacity of the exposure-response curve. These are derived from deviance statistics for hierarchical models. Following McCullagh and Nelder (1989) and Venables and Ripley (1994), this type of analysis is called the analysis of deviance. The analysis of deviance statistics and R^2 statistic are in a table in the output file. The command `deviance.fit()` will also cause them to be calculated and written to the output file.

After running CatReg, the fitted curve is available for further analysis. As an example of the analysis of deviance, consider the output file in Table 5-5. The summary output in the table, which shows the coefficient estimates, indicates that seven parameters have been estimated. There are 64 treatment groups (2 species x 2 targets x 4 concentrations x 4 durations), so total degrees of freedom (df) is 64 x (3 severity levels - 1) = 128 (unadjusted for SEV1 and SEV2). The estimates of the SEV1 and SEV2 intercepts provide no information about how well the curve fits the data, so it is customary to adjust the total degrees of freedom for them in the analysis of deviance table, leaving 128 - 2 = 126 degrees of freedom. The model has 5 parameters, aside from SEV1 and SEV2, which are MU:L:INTERCEPT, RT:C:INTERCEPT,

RT:L:INTERCEPT, LG10C0NC, AND LG10TIME. This leaves 126 – 5 = 121 degrees of freedom for the residual deviance.

The analysis of deviance table for the example shown in Table 5-5 partitions the total deviance into the sum of two components, the "model" deviance and the "residual" deviance. The total deviance is the deviance when the only parameters in the exposure-response curve (or "model" in the terminology being used here to show the comparison with the analysis of variance) are the intercepts, SEV1 and SEV2. This is referred to as the null model because it contains no explanatory variables. The residual deviance is simply the deviance of the fitted model, which includes the 5 parameters, aside from SEV1 and SEV2, i.e., MU:L:INTERCEPT, RT:C:INTERCEPT, RT:L:INTERCEPT, LG10C0NC, and LG10TIME, as explanatory variables. The model deviance is that part of the total deviance that is explained by the model (the proportion is the generalized R^2). In this example, 46.4% of the total deviance is explained by the model. R^2 is a general measure of the proportion of the variation in the response that is accounted for by the explanatory variables.

The mean deviance entries, labeled as "Mean.Dev," are computed as deviance divided by degrees of freedom. An approximate F-test of the model is obtained as the ratio of model to residual mean deviations. This is an approximate F-statistic in large samples under the ordinal regression model with independent responses. This statistic tests the null hypothesis that all explanatory variables can be dropped from the model. The same hypothesis may be tested using partest. The two results will be similar if the responses are independent, and the residual degrees of freedom are reasonably large, say larger than 15.

Generally, the F-test will reject the null hypothesis unless the sample size is very small or the model fits poorly. It merely verifies that there is *some* relationship between the response and the explanatory variables. The generalized R^2 statistic often will be of more direct interest as a measure of the explanatory value of the variables in the model.

Cluster sampling invalidates the large sample F distribution of the generalized F-statistic. However, it is common practice to compute F as a rough guideline; see Venables and Ripley (1994, p. 187). In any case, the R^2 statistic gives an idea of how much variation in the response is accounted for by the explanatory variables.

8. MAKING PLOTS

8.1. The Plots and Options

CatReg includes seven functions for making plots after an exposure-response curve has been fit to the data (commands to create plots are input at **###(6)Issue Commands ###** in Table 5-1). The seven plots are listed with brief descriptions in Table 8-1.

TABLE 8-1. PLOT FUNCTIONS AVAILABLE IN CATREG.

catplot() Concentration on y-axis, duration on x-axis. Plot of ERC line with confidence interval for a single severity level and stratum, and response data for all severity levels.

stratplot() Concentration on y-axis, duration on x-axis. Plot of ERC lines and response data for all strata for a single severity level.

confplot() Strata for a single severity level displayed on y-axis, concentration on x-axis. Plot of ERC with confidence interval for the unstratified model, and for individual strata of the model, for a single exposure duration.

prplot() Probability of exceeding a specified severity level on y-axis, concentration or duration on x-axis. Plots the probability curve as a function of concentration, keeping duration fixed, or as a function of duration with concentration fixed.

dataplot() Concentration on y-axis, duration on x-axis. Plots response data for all severity levels combined by stratum.

allsevsplot() Concentration on y-axis, duration on x-axis. Plots ERC lines and response data for all severity levels for a single stratum.

devplot() Generalized deviance residuals on y-axis and data observation number, log-concentration, or log-duration on x-axis.

Several options available for use with various plots are listed in Table 8-2. An option is implemented by inserting it as an argument in the parentheses at the end of the name of the plot function. **replace=T** applies to all plot commands. It sends the new plot to the currently active graphics window instead of generating a new one and overwrites the previous graph. For example, the command catplot(replace=T) would plot in the currently active window and replace the previous graph. **cl=** applies to catplot and confplot. It is used to change the *two-sided* confidence limit used in a plot from the 90% default (cl=0.9).

TABLE 8-2. PLOT OPTIONS.

Option	catplot	stratplot	confplot	prplot	dataplot	allsevsplot	devplot
replace=T	X	X	X	X	X	X	X
cl=	X		X				
log="x"	X	X	X		X	X	
log="y"	X	X			X	X	
log="xy"	X	X			X	X	
log=" "	X	X		X	X	X	
Xlim=	X	X	X	X	X	X	
Ylim=	X	X		X	X	X	
jitter=T	X	X		X	X	X	
title=	X	X	X	X	X	X	
time=				X			
conc=				X			
prcurve=F				X			

For example, the command catplot(cl=0.95)would produce a 95% two-sided confidence interval instead of the default 90%. **log="y", log="x", log="xy", and log= ""** mean use a log scale (base 10) on the y-axis only, the x-axis only, both axes, neither axis, respectively. The defaults are log="y" for catplot, stratplot, dataplot, and allsevsplot, log="x" for prplot, and log = "" for confplot. **xlim=** and **ylim=** are used to set the ranges of x-axis and y-axis, respectively, which are determined from the data by default. They take the form xlim=c(lo,hi) and ylim=c(lo, hi), where lo is the lower limit, and hi is the upper limit, e.g., catplot(ylim=c(.1,100000)) specifies that the concentration axis (y-axis) runs from 0.1 to 100,000. **jitter=T** adds small random offsets to coincident observations. Although the amount added is bounded, the randomness of the offset causes each jitter of the data to produce a slightly different plot. Jittering is useful to get a sense of data density in cases where observations pile on top of one another. **title=** allows the user to create the title of a plot in place of the default. For example, catplot(title = "My own title"). **time =** and **conc =** set the duration or concentration level to be used in prplot, e.g., prplot(time = 2)or prplot(conc = 1000). **prcurve = F** suppreses the curve from being included in the prplot, e.g., prplot(time = 2, prcurve=F).

　　　　Several plotting commands produce a pop-up menu requesting whether a legend is desired. If "yes" is selected, use the mouse to position the upper left corner of the legend on the graph and then left-click the mouse. A graph may be saved, copied, or printed in two ways. First

58

either left-click on the graph or left-click on "Windows" and select the graph from the menu. The user can then either right-click the mouse on the graph to produce a menu, or left-click on "File" in the R commands. Either way will allow the graph to be saved as a Metafile or Postscript; the latter way (using the "File" menu) provides four additional file types (PDF, Png, Bmp, and Jpeg). On exiting R, the user also is offered the choice of saving all graphs.

As discussed in Section 5.1, after fitting an exposure-response curve to the data, CatReg calculates extra risk concentrations for an ERC percentile, severity level, and (one-sided) confidence interval percentile determined by the user. Using Table 5-1 as an example, this occurs in the section at **`###(5)Calculate and output extra risk ###.`** Specifically, the queries for the ERC settings are:

```
Percentile for ERC? (default=10):
Severity level for ERC? (default = 1):
Percentile for 1-sided upper and lower confidence intervals?
(default=95):
```

The user either accepts the default values shown in parentheses or inputs different values to use. The ERC settings remain current until they are subsequently changed by execution of the function `ecdata`.

The function `ecdata` is described in Section 9.1. Briefly, the user enters the command `ecdata()` and then responds to the same queries as above (the last query, whether to write ERC data to files is not shown). For example (user input is in boldface),

>ecdata()
```
Percentile for ERC? (default=10):5
Severity level for ERC? (default = 1):2
Percentile for 1-sided upper and lower confidence intervals?
(default=95):85
```

would change the current ERC settings to 5, 2, 85. The plotting functions that require a severity level and/or percentile for the ERC use the values in the current ERC settings. The two functions

that plot confidence intervals for the ERC, however, namely `catplot` and `confplot`, *do not* use the current setting for percentile of confidence intervals. They both use a default level of 95% for one-sided upper and lower confidence intervals (equivalently, 90% two-sided intervals). In both cases, the default can be overridden by use of the `cl=` option discussed above. If the current ERC settings need to be changed for the sake of making a plot, just execute `ecdata` prior to making the plot.

The various plots are described and illustrated by example in the next section. All of the examples were produced from commands after running CatReg on the input file chemx.csv, with the setup options as indicated in the output file in Table 5-5. The default options for the ERC settings (10, 1, 95, for the ERC percentile, ERC severity level, and ERC percentile for one-sided confidence intervals, respectively) were used. They were the current ERC settings when the plots were made. Pop-up menus give the user a choice of strata, for the ERC severity level used, and the option to include a legend, as appropriate to the specific plot and whether there is stratification.

8.2.   Concentrations and Durations for Designated ERC and Severity (catplot)

The command `catplot()` plots the response data and graphs the ERC with confidence interval, for a single severity level and stratum. Exposure concentration is on the y-axis and exposure duration on the x-axis. The current ERC settings are used for ERC percentile and severity level. The confidence interval percentile is 90% (two-sided) but may be overridden by the `cl =` option. The stratum to use and whether to include a legend are determined from pop-up menus. By default, concentration and duration are graphed on a log-linear scale. This type of graph is useful for showing how extra risk changes with concentration or duration. Be aware that the weighting of individual points is not shown. The number of points in each severity category will be displayed in the R command window, along with the number of hidden points. Each call to `catplot` generates a new graphics window. Thus, repeated use of `catplot` allows comparison of results across strata. The options available for use with `catplot` are in Table 8-2.

Figure 8-1 was produced with the instruction `catplot(jitter=T)`. The stratum for the ERC line was selected from a pop-up menu giving four choices: SEV1:MU:C, SEV1:MU:L, SEV1:RT:C, SEV1:RT:L (the choices were determined from the summary table of estimates in
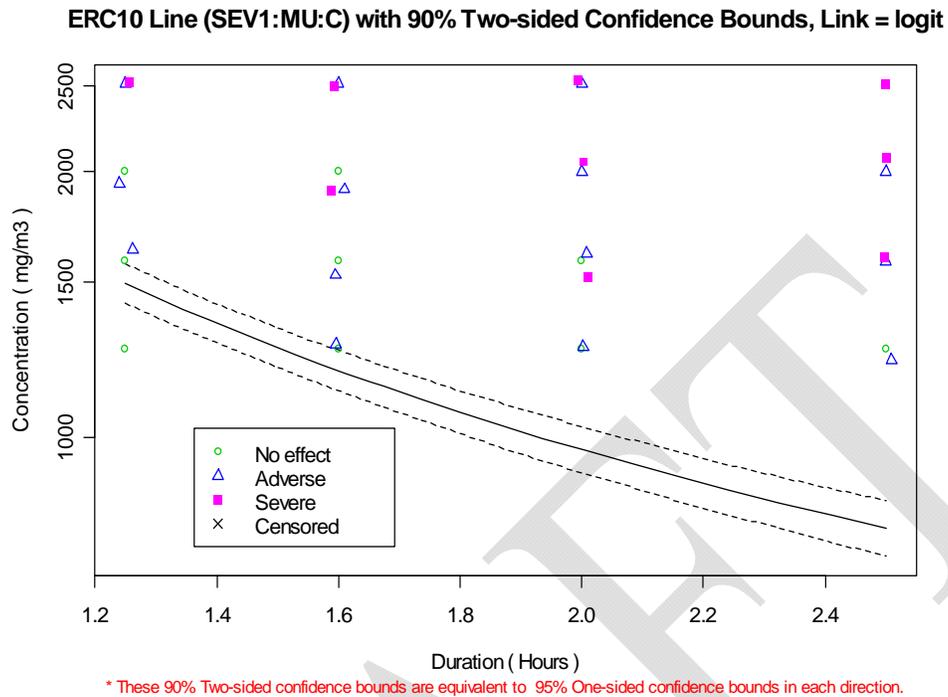
Figure 8-1. Illustration of plot function `catplot`.

the output file in Table 5-5). The choice to include a legend was selected from a second menu.
The x-y locations of symbols on the graph indicate the exposure concentrations and durations of
observations on mice with the central nervous system as the target organ. The symbol itself
indicates the severity category, as shown in the legend. The lines on the graph are the estimated
ERC10 (solid line) and upper and lower 95% one-sided confidence bounds (dashed lines)
(equivalently, two sided 90% confidence bounds). Slicing the graph vertically at a specific
duration gives the confidence interval that `confplot` would graph for that duration. As
expected, longer durations require lower concentrations to achieve the estimated 10% level of
extra risk.

**8.3   Concentrations and Durations for Designated Probability and Severity
of Strata (stratplot)**

The command `stratplot()` plots the response data by stratum for all severity levels
and graphs the ERC (without confidence interval) by stratum for a single severity level.

Exposure concentration is on the y-axis and exposure duration on the x-axis. The current ERC settings are used (for the ERC percentile and ERC severity level). Response to a pop-up menu determines whether to include a legend. By default, concentration and duration are graphed on a log-linear scale. Be aware that the weighting of individual points is not shown. The number of points in each severity category will be displayed in the R command window, along with the number of hidden points. The options available for use with `stratplot` are listed in Table 8-2. This graph provides comparison of the ERC curves for different strata by plotting them on the same graph. Figure 8-2 was produced by entering the command `stratplot(jitter=T)`.

8.4. Extra Risk Concentration at Desired Duration (confplot)

The command `confplot(time= )` displays the ERC with confidence interval for the unstratified model and for individual strata of the stratified model, for a single severity level at the exposure duration set in the argument `time= `, e.g., `confplot(time=1.5)` for 1.5 hours exposure. The time argument is required unless duration is not included in the data as an explanatory variable. The current ERC settings are used for the ERC percentile and severity level. The confidence interval percentile is 90% (two-sided) but may be overridden by using the `cl=` option. This graph is useful for comparing ERC estimates and confidence intervals among strata, and comparing individual strata with the unstratified model.

Figure 8-3 was produced from the command `confplot(time=2)`. Concentration is on the x-axis. The central dot for each stratum is the estimated ERC10 for severity level 1. The vertical solid line is the estimated ERC10 for the unstratified exposure-response curve, and the vertical dashed lines are the associated confidence intervals (90% for two-sided bounds, determined from the ERC settings since the option `cl = ` is not used here). To display the confidence interval for another duration, repeat the `confplot` command with the option `time= ` reset to the new duration. Execute `ecdata` to change the ERC percentile or the severity level and the option `cl = ` to modify the percentile of the confidence intervals.

8.5. Probability Versus One Explanatory Variable (prplot)

The command `prplot()` displays the probability curve (i.e., the exposure-response curve) as a function of concentration, keeping duration fixed, or as a function of duration with
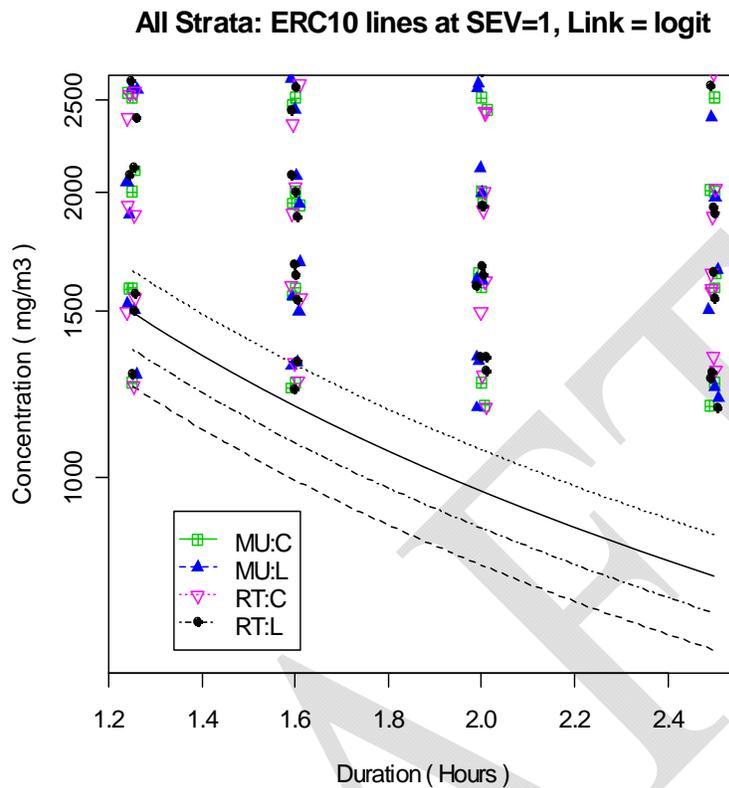
Figure 8-2. Illustration of plot function `stratplot`.

concentration fixed. If either of these variables is constant in the data, then `prplot` graphs the probability against the nonconstant variable. If both concentration and duration vary in the data dataset, then the user needs to tell `prplot` which variable to hold constant. This is done by specifying either `time=` or `conc=` as an argument. For example, to plot the response probability versus duration at a fixed concentration, use `prplot(conc= )`, such as `prplot(conc = 2000)`. To plot the response probability versus concentration at a fixed duration, use `prplot(time= )`, e.g., `prplot(time = 2)`.

The severity level is determined by the current ERC settings. The choice of stratum and whether to include a legend are determined by response to pop-up menus. The R command window will show the number of data points at the required duration or concentration, and the graph will show the probability of exceeding a particular severity level, for a specified stratum (when there is stratification) versus either concentration or time. The available options are listed
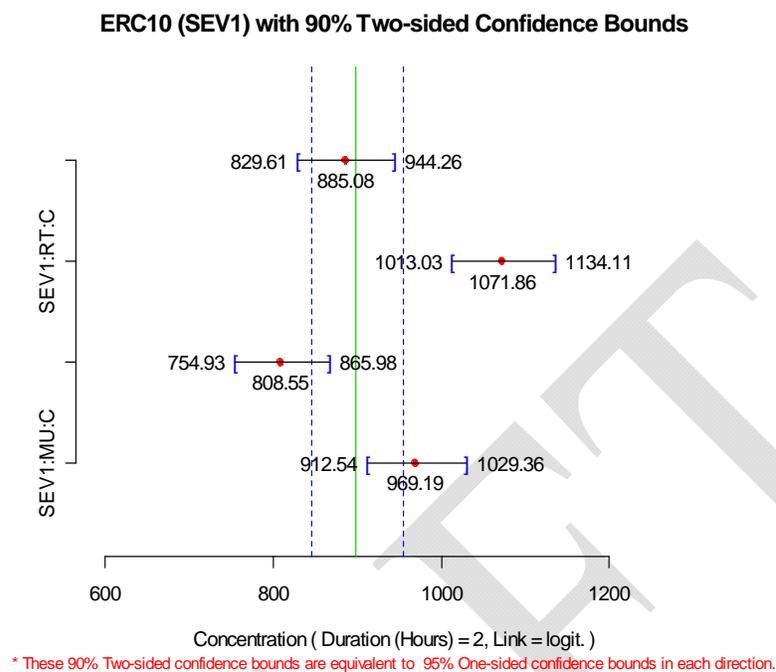
ERC10 (SEV1) with 90% Two-sided Confidence Bounds

Concentration (Duration (Hours) = 2, Link = logit.)

* These 90% Two-sided confidence bounds are equivalent to 95% One-sided confidence bounds in each direction.

Figure 8-3. Illustration of plot function `confplot`.

in Table 8-2. Note that the option `prcurve=F` suppresses the probability curve, leaving just the data points in the plot. The function `prplot` is useful to help assess whether the fitted probability curve is consistent with the data, and for representing the risk over a range of exposure levels.

Figure 8-4 was produced with the command `prplot(time=1.25).` The curve in Figure 8-4 plots the probability of the occurrence of a severity level 1 response or greater for the liver target organ in mice (species MU) as a function of concentration, with duration fixed at 1.25 h. The value used for time need not be an exposure time in the data set. When that occurs, the probability curve is displayed but there are no data to display.

In these graphs, the vertical location of a symbol represents whether the response at a particular concentration was equal to or greater than the severity of interest (adverse, severity 1, for Figure 8-4). No effect responses are plotted at Pr = 0. Adverse effects and severe effects are both plotted at Pr = 1. There are no censored observations. If the input file had contained some observations censored as [0,1], they would have been graphed as "×" at Y = 0.5 on the severity
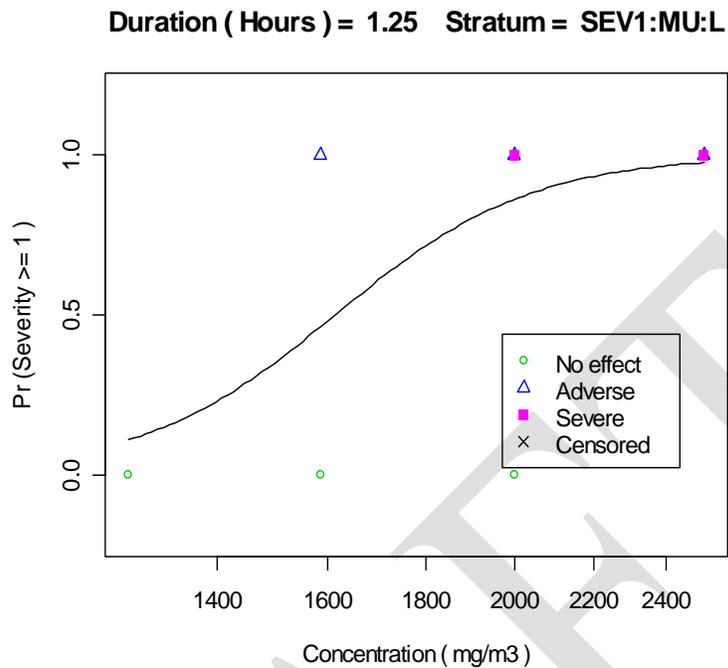
Duration (Hours) = 1.25   Stratum = SEV1:MU:L

Figure 8-4. Illustration of plot function `prplot`.

≥ 1 graph, and as "×" at $Y = 0$ on the severity ≥ 2 graph. This is because, in the first case, it is not known whether the censored observation meets the severity threshold (i.e., whether severity ≥ 1), whereas in the second case the severity is known to be less than 2.

8.6.   Data Plotted by Stratum (dataplot)

The command `dataplot()` plots the response data for all severity levels by stratum, without a response probability curve, as shown in Figure 8-5. Exposure concentration is on the y-axis and exposure duration on the x-axis. Notice that Figure 8-5 is very similar to Figure 8-2, except that Figure 8-5 does not show the ERC10 lines. After asking whether a legend is desired, the command window shows the number of points in each stratum, the total number of points for the analysis, and the number of hidden points.
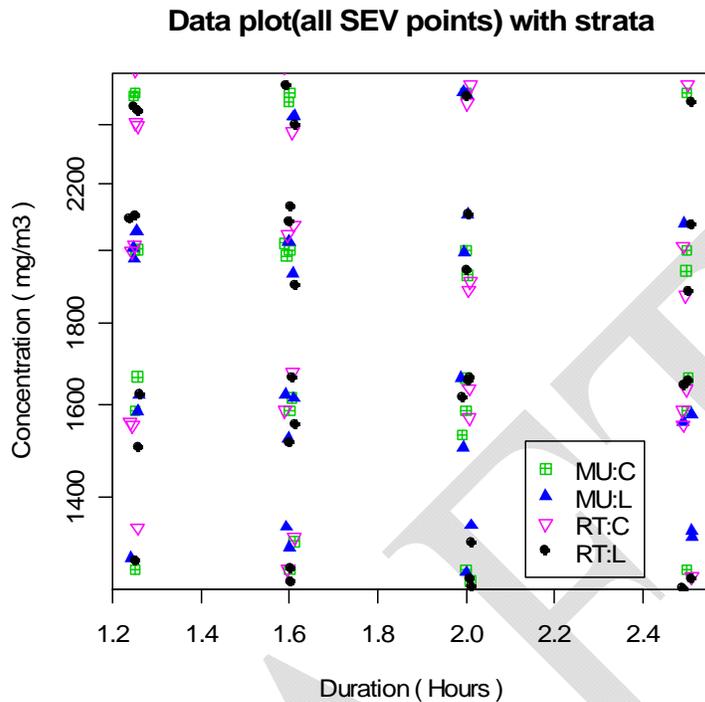
Data plot(all SEV points) with strata

Figure 8-5. Illustration of plot function `dataplot`.

8.7.   Data and ERC Lines for Single Stratum, all Severity Levels (allsevsplot)

The command `allsevsplot()` plots the response data and ERC lines for a single stratum at all severity levels. The stratum is selected by the user from a pop-up menu. A second pop-up menu provides an option to include a legend. The points and ERC lines are distinguished by severity level. Figure 8-6 was produced with the command `allsevsplot(jitter=T)`, then choosing stratum MU:C and electing to include a legend. This plot is useful for comparing the ERC across severity levels.
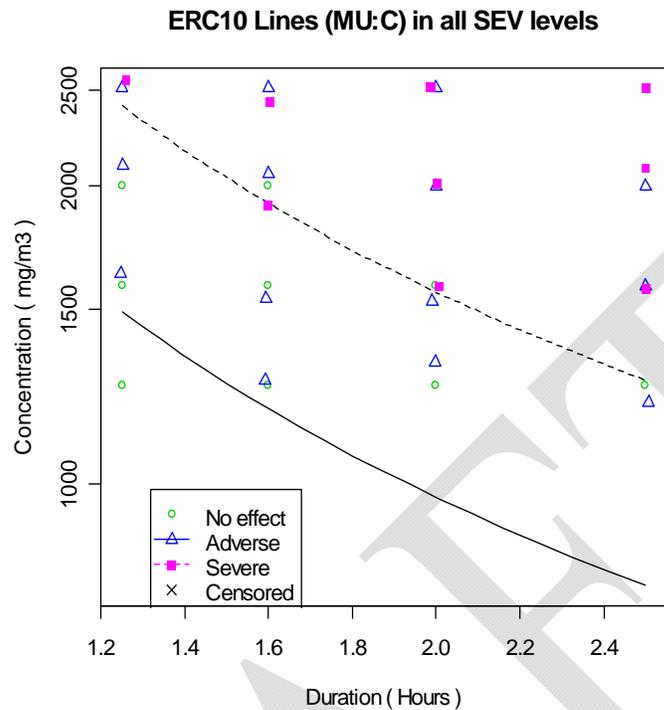
Figure 8-6. Illustration of plot function `allsevsplot`.

8.8.   Contribution to Deviance for Individual Datum (devplot)

How well are individual observations explained by the exposure-response curve? In linear regression, residual plots help to address that question, that is, plots of the differences between observed and predicted responses. Here the response is binary or ordinal, so there is a scaling problem in defining a " residual." Various possibilities exist. CatReg uses the individual components of the deviance statistic to measure how well individual observations are explained by the exposure-response curve. Using deviance residuals is common practice in generalized linear modeling (see Venables and Ripley [1994, p. 188-189]). For interval-censored ordinal data, other types of residuals are undefined, whereas the deviance residuals are defined in the usual way. The observations that contribute to any lack of fit of the exposure-response curve can be identified by examination of the individual contributions to the deviance.

The command `devplot()` produces a diagnostic plot of generalized deviance residuals versus observation number (excluding filtered data), concentration (or log-concentration), or duration (or log-duration), depending on the user selection from a pop-up menu. If the model is stratified, the points are labeled by strata. A pop-up menu makes inclusion of a legend an option.

Generally, plotting deviances versus observation number (Obs.Number) is a good choice. It provides a representation of the relative effectiveness of the model in fitting the different observations or strata. Devplot does not require jittering because each observation is represented uniquely under the Obs.Number option. The fit is suspect if one or a few observations have much larger deviance residuals than the remaining observations because the fit may be unduly influenced by these observations. If one stratum has large deviances, the model may be inadequate for this stratum. Rerunning the model without this stratum would allow one to determine whether the results for the other groups are heavily influenced by the poorly fit subset.

Plots of deviance versus concentration (or log-concentration) or time (or log-time) are useful in studying the adequacy of the functional form of the regression relationship. Trends in the deviances would suggest a problem with the functional form. One should be aware that differences in the density of the data at different concentrations or durations will affect the perception. Regions of the plot with more data will tend to have more spread in the deviances because of random variation, even if the model is adequate for the data. Figure 8-7 was produced by entering devplot() and then selecting Obs.Number from the pop-up menu.

Figure 8-8 is an example of devplot for a different input file. The input and output files are not shown, but the same targets (C for central nervous system and L for liver) apply, and HU (human) is added as a species in addition to MU (mice) and RT (rats). Intercepts were stratified on "Species" and "Target". The notable feature of the figure is that the liver data for mice are relatively poorly described by the curve (i.e., residual deviances are relatively high) in comparison to most of the other data. These data may or may not be influential on the fitted exposure-response curve. To examine whether they are, the curve could be refit after filtering out liver data for mice.

8.9.   Delete All Plots (rmplots)

The rmplots() command is particularly useful for keeping the number of graphics windows under control when fitting several CatReg models in one R session. It removes all graphics windows created since the beginning of the R session or since the last call to rmplots. Enter the command rmplots() to close all currently active graphics windows without exiting

R. It has no effect on the command window or on subsequent use of the graphical display functions.

8.10.   Printing Graphics

There are several ways to print a graph: (1) right-click on the graph and left-click on "Print", or (2) either left-click on the graph or left-click on "Windows" and then left-click the appropriate choice from the Windows menu, then left-click on "File", and left-click on "Print".
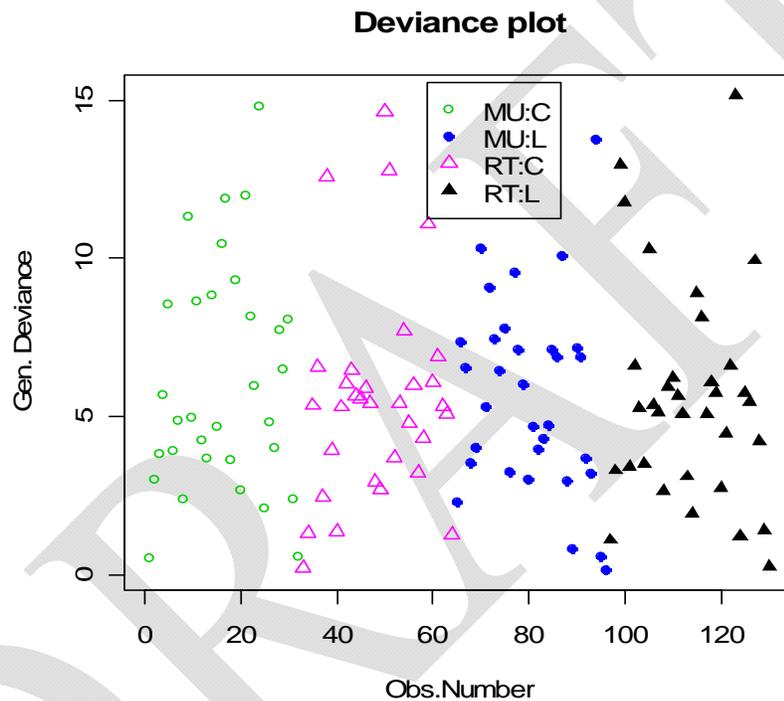
Deviance plot

Figure 8-7. Illustration of plot function `devplot` - Example 1.
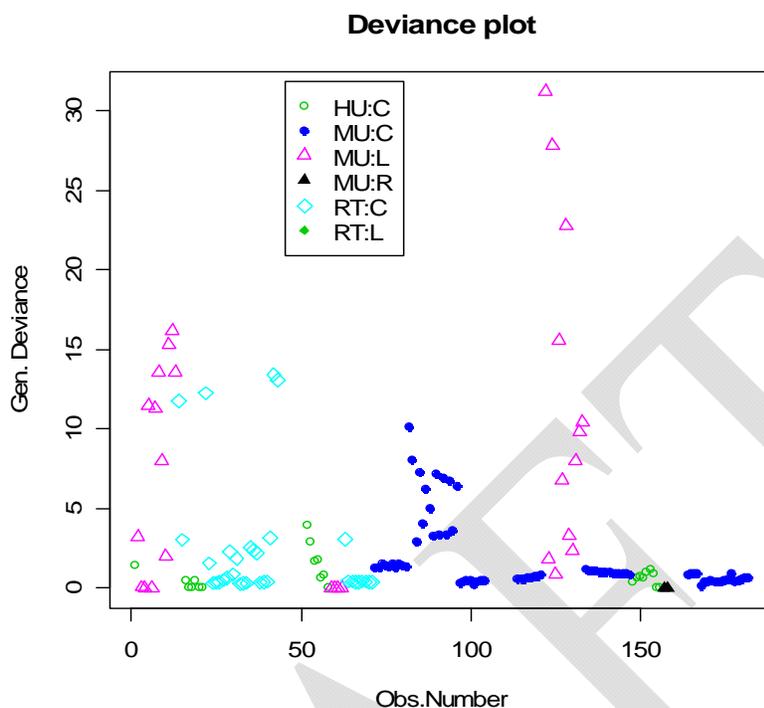
Deviance plot

Figure 8-8. Illustration of plot function `devplot` - Example 2

9. OTHER CatReg FUNCTIONS

9.1. Change Probability and Severity Category (ecdata)

As discussed in Section 5.1, CatReg calculates extra risk concentrations once the model estimation step is complete (at location **### (5) Calculate and output extra risk ###** in Table 5-1). It queries the user for ERC settings (ERC percentile, ERC severity level, and percentile for one-sided upper and lower confidence intervals) to be used in ERC calculations and asks the user whether to write the ERC results to text files. Using the command `ecdata()` (at location **###(6)Issue Commands ###** in Table 5-1) just repeats those queries so that the user can change the ERC settings for subsequent use in making plots or to repeat calculations of ERC estimates at a different ERC percentile, severity level, or confidence interval percentile.

If the user responds with "y" to the query `Write ERC data to text files?(n),` then a separate text file is created for each stratum with the same name as the output file requested by the user in the current run of CatReg (e.g., chemx) except with the file names appended numerically, i.e., with 1, 2, …, depending on how many strata there are (e.g.,

70

chemx1, chemx2, …). Note that existing files with those names will be over-written. For each stratum, `ecdata` makes the ERC calculations for 20 exposure durations (times), equally-spaced over the range of exposure durations in the input file, including both the smallest and largest times. The results are stored in the R object `ECline` and the instruction `>ECline` will display the results on the screen for each stratum. The graphical functions `catplot` and `stratplot` access `ECline`.

An example of the use of `ecdata()`with chemx.csv, for output file as shown in Table 5-5, follows (user input is in boldface).

```
> ecdata()
Percentile for ERC? (default=10): 5
Severity level for ERC? (default = 1): 2
Percentile for 1-sided upper and lower confidence intervals?
(default=95): 90
ERC data computed...
Write ERC data to text files? (n):y
Writing ERC data to ascii files...
Done - see chemx.out for file names.
```

The ERC5 values were written to four text files (chemx1, chemx2, chemx3, and chemx4) corresponding to the four strata in the analysis. The output file chemx1 is displayed Table 9-1.

9.2.   Extra Risk Concentration for Desired Durations (ec3table)

To obtain extra risk concentration estimates for user-specified durations without reading them from a graph, use `ec3table(duration = time)`, with time expressed in hours (even if the log-transform has been used for time in the run of CatReg), e.g., `ec3table(duration = 2)`. More than one exposure duration can be entered by using the format `ec3table(duration = c(time1, time2, time3,etc.))`, where time1, time2, …., are the times of interest, e.g., `ec3table(duration = c(1.5,2.0,2.5))`. Any number of durations can be requested.   A table of the ERC values and confidence intervals for the exposure durations requested is produced and displayed on the screen, by stratum, for a

specified severity level. To direct the table of values to the output file instead of to the screen , include the argument `printout = T.` The ERC percentile, the percentile for the one-sided upper and lower confidence bounds, and the severity level are taken from the current ERC settings.

For example, during the run of CatReg that produced the output in Table 5-5, the ERC percentile was 10, the severity level was 1, and the confidence percentile was 95 (apparent from the ERC summary at the end of Table 5-5). To obtain the numerical estimates of the ERC10 at the observed durations, for severity level 1, with 95% one-sided confidence limits, we would use the instruction `ec3table(duration = c(1.25,1.6, 2, 2.5))`. To reset the ERC percentile, severity level, or confidence limit percentile, without executing CatReg again, execute the `ecdata` function prior to executing `ec3table`. The following example changes the ERC percentile to 5 (user input is in boldface):

TABLE 9-1. FILE CHEMX1 OF ERC5 VALUES FOR STRATUM SEV2:MU:C PRODUCED BY COMMAND `ECDATA().`

Time	ERC5	sd	LB90	UB90
1.25	2208.16	74.35	2114.91	2305.53
1.3	2135.2	70	2047.35	2226.81
1.34	2064.64	66.05	1981.7	2151.05
1.39	1996.41	62.5	1917.9	2078.14
1.45	1930.44	59.32	1855.9	2007.98
1.5	1866.65	56.5	1795.63	1940.49
1.56	1804.97	54.03	1737.04	1875.55
1.61	1745.32	51.87	1680.09	1813.09
1.67	1687.65	50.02	1624.75	1752.99
1.74	1631.88	48.45	1570.96	1695.17
1.8	1577.96	47.13	1518.7	1639.52
1.87	1525.81	46.04	1467.94	1585.97
1.94	1475.39	45.15	418.66	1534.4
2.01	1426.64	44.44	1370.82	1484.74
2.08	1379.5	43.88	1324.4	1436.89
2.16	1333.91	43.45	1279.38	1390.77
2.24	1289.83	43.12	1235.74	1346.3
2.32	1247.21	42.89	1193.44	1303.4
2.41	1206	42.72	1152.47	1262.01
2.5	1166.15	42.61	1112.8	1222.05

```
 ** Stratum :SEV2:MU:C
 ** Risk Type :extra risk
 ** Above UB90 and LB90:
    These one-sided 90% lower bound and one-sided 90% upper bound confidence
intervals are equivalent to the lower bound and upper bound of two-sided 80%
confidence intervals.
```

```
> ecdata()
Percentile for ERC? (default=10): 5
Severity level for ERC? (default = 1): 1
Percentile for 1-sided upper and lower confidence intervals?
(default=95): 95
ERC data computed...
Write ERC data to text files? (n):
```

The command ec3table(duration = c(1.25,1.6, 2, 2.5))will then produce the calculation and screen display of ERC5 estimates, with 95% confidence bounds, for severity level 1, as displayed in Table 9-2.

TABLE 9-2. MONITOR DISPLAY OF ERC5 ESTIMATES PRODUCED BY COMMAND EC3TABLE(DURATION = C(1.25,1.6, 2, 2.5)).

```
< ERC5 at specific time point(s) >
 : One-sided 95% Lower bound and One-sided 95% Upper bound

* Risk Type : extra risk

*  ERC5 at SEV1:MU:C
                ERC5 Lower Bound of ERC5 Upper Bound of ERC5
1.25 Hours 1368.8089           1292.5118           1449.6099
1.60 Hours 1090.4150           1027.7983           1156.8464
2.00 Hours  887.8264            830.1670            949.4905
2.50 Hours  722.8769            667.7075            782.6046


*  ERC5 at SEV1:MU:L
                ERC5 Lower Bound of ERC5 Upper Bound of ERC5
1.25 Hours 1141.9330           1073.5657           1214.6542
1.60 Hours  909.6820            851.4345            971.9143
2.00 Hours  740.6719            686.7422            798.8367
2.50 Hours  603.0622            551.9808            658.8708


*  ERC5 at SEV1:RT:C
                ERC5 Lower Bound of ERC5 Upper Bound of ERC5
1.25 Hours 1513.8125           1431.1282           1601.2739
1.60 Hours 1205.9271           1139.9345           1275.7402
2.00 Hours  981.8775            921.6875           1045.9982
2.50 Hours  799.4542            741.7205            861.6817

*  ERC5 at SEV1:RT:L
                ERC5 Lower Bound of ERC5 Upper Bound of ERC5
1.25 Hours 1250.0238           1177.8598           1326.6091
1.60 Hours  995.7888            935.1591           1060.3494
2.00 Hours  810.7809            754.6849            871.0465
2.50 Hours  660.1456            606.7391            718.2531

These one-sided 95% lower bound and one-sided 95% upper bound
confidence intervals are equivalent to the lower bound and upper bound of
two-sided 90% confidence intervals
```

9.3.   Summary for Most Recent Analysis (catreg.summary)

Entering catreg.summary()provides summary information that is displayed on the screen. Included are the input file name, information on the setup options, and results for the fitted exposure-response curve, such as parameter estimates, standard errors, and Z-statistics for

testing individual parameters for removal from the model. This is the same data summary displayed on the screen after optimization is complete (in the section beginning with **`###(3)Display and output summary results ###`** in Table 5-1) and in the output file. The example in Table 9-3 applies to running CatReg using input file chemx.csv, with output file displayed in Table 5-5.

TABLE 9-3. MONITOR DISPLAY OF CATREG SUMMARY PORTION OF AN OUTPUT FILE PRODUCED BY COMMAND `CATREG.SUMMARY()`.

```
Input file   : chemx.csv
Filtered data: none
Model        : cumulative odds model
Link         : logit
Clustering   : none
Message      :
Iterations   : 41 14
Deviance     : 741.0266
Residual DF  : 121
AIC          : 755.0266

Scale:
    Concentration: log10( mg/m3 )
    Duration     : log10( Hours )

Stratification:
    Intercept    : Species Target
    Concentration:
    Duration     :

Coefficients:
            Estimate        Std. Error   Z-Test=0      p-value
SEV1
            -66.2375249     4.2947663    -15.422847    0.00001
SEV2
            -70.3128027     4.4629521    -5.754774     0.00001
MU:C:INTERCEPT
            0.0000000       0.0000000    NA            NA
MU:L:INTERCEPT
            1.5443003       0.2803179    5.509103      0.00001
RT:C:INTERCEPT
            -0.8580588      0.2701988    -3.175658     0.00149
RT:L:INTERCEPT
            0.7735915       0.2703681    2.861253      0.00422
LG10CONC
            19.6220725      1.2858838    15.259600     0.00001
LG10TIME
            18.0738765      1.2265874    14.735091     0.00001
```

9.4. Information on Designated Attributes (attributes(fits))

To access components of the fitted model that are not shown in the CatReg summary or text files, enter attributes(fits) to see what components are available; a list will be displayed.

```
> attributes(fits)
$names:
 [1] "deviance"   "df.residual" "message"   "coefficients"
     "gradient"   "hessian"     "sand.var"
 [8] "itr.num"    "f.evals"     "g.evals"   "scale"
     "link"       "model"       "xnames"
[15] "xnames.gp"  "nsevcat"     "ref.cell"
```

To see a particular component, type the object name `fits`, and the component name joined by "$". For example,

```
> fits$deviance
[1] 435.8275
```

reveals that 435.8275 is the deviance for the model.

10.   UTILITIES

CatReg includes a number of utility functions for such objectives as renaming required input variables, viewing input data, converting data files to comma-separated files, combining severity categories, recoding variables, and recoding missing values.

10.1.   Renaming Required Input Variables

CatReg requires certain variables in order to perform an analysis. The default names for these fields are contained in an R object named `required`. Often the variable names in a data set are different from the default names. To display the current default variable names, enter `required` at the R prompt. The current names are then shown for the six types of required variables, as shown below.

```
> required
 conc      time     loscore  hiscore   incid    nsub
"mg/m3"   "Hours"  "SevLo"  "SevHi"   "Incid"  "Nsub"
```

Note that the variable names are displayed in quotes, whereas the names of the data fields are aligned above them. If the data set variable names are different from these, use the CatReg function `rename` to replace the default names with the variable names in the data set.

One application for renaming variables is to make a comparison of results using censored data to those using the best judgment scores (i.e., `BestNum`). This type of sensitivity analysis can be a useful indication of the impact of uncertain severity scores on the outcome. To use `BestNum`, the default names `SevLo` and `SevHi` must be changed to `BestNum`. CatReg will no longer see the data entries as censored and will use the severity category in the `BestNum` column for both `SevLo` and `SevHi`. This can be accomplished with the following commands.

```
> rename(required)->required
**** Change required variable names ****
 Current required variables:
 mg/m3 Hours SevLo SevHi Incid Nsub
 Rename which variables? (none):
1: SevLo SevHi
3:
 Current names:
 SevLo SevHi
 Enter new names:
1: BestNum BestNum
```

And to confirm the change:
```
> required
 conc      time     loscore     hiscore    incid     nsub
"mg/m3"   "Hours"  "Bestnum"   "Bestnum"  "Incid"   "Nsub"
```

To avoid an ambiguity error (having two variables named "BestNum") when changing the variables back to SevLo and SevHi, use the following commands.

```
> required["loscore"]<-"SevLo"
> required["hiscore"]<-"SevHi"
```

77

The composite symbol "`->`" is the R assignment symbol. It is composed of the "`-`" and "`>`" symbols. Assignments may be forwards as in the example, or backwards as "`<-`". The object `required` was replaced by a modified version of itself, entering the new variable names in response to the query. The change can be verified by displaying the required variable names again:

```
> required
 conc     time     loscore  hiscore   incid    nsub
"mg/m3"  "Hours"  "SevLo"  "SevHi"   "Incid" "Nsub"
```

The data fields `conc`, `loscore`, `incid`, and `nsub` are required to be in the input data set. The remaining two fields are set to the following default values if they are not found in the input data: `time=1`, `hiscore=loscore`. To reset the names of the required variables to the CatReg default names, omit the argument of the rename function: `> rename() ->` `required` .

10.2. Viewing Input Data

To view the contents of a comma-separated data file during an R session, use the CatReg `view` utility. Briefly, the function called `view(infile="`*filename*`")` will scan the named file, display the first record (usually a list of variable names), and ask which variables to display. The default is to display all variables. A subset of data records may be specified using the `row=` option (see the example below).

```
> view(infile="chemx.csv", row=1:5)
Read 1179 items
 Variables:
 Exp. Group Species Target mg/m3 Hours SevLo Nsub Incid
  View which variables? (all):
1:
Read 0 items
     Exp. Group Species Target mg/m3  Hours  SevLo Nsub Incid
```

```
[1,] "1"  "1"  "MU"   "C"    "1259" "1.25" "0"   "10" "10"
[2,] "1"  "2"  "MU"   "C"    "1259" "1.6"  "0"   "10" "9"
[3,] "1"  "2"  "MU"   "C"    "1259" "1.6"  "1"   "10" "1"
[4,] "1"  "3"  "MU"   "C"    "1259" "2"    "0"   "10" "4"
[5,] "1"  "3"  "MU"   "C"    "1259" "2"    "1"   "10" "6"
```

If the `row=` option is omitted, then all records will be displayed. To load the data into an R matrix use the R assignment `->`, e.g., `view(infile="chemx.csv") -> chemx`, which loads the sample data, `chemx.csv`, into an R matrix named `chemx`. The data are stored as a character matrix. Before applying any R statistical functions to the data, it is necessary to convert quantitative variables from characters to numeric values. A facile way to do this is to convert to an R data frame, as in the following example: `data.frame(chemx) -> chemx`. This replaces the character matrix `chemx` with a data frame of the same name. Typing `chemx` at the R prompt will display the contents of this new data frame. A data frame is a matrix of data whose columns may be either character strings or numeric values. Many R functions operate on data frames.

10.3. Converting Data Files to Comma-Separated Files

CatReg requires the input data to be comma-separated. If the input data file is created using Excel, it can be saved as a csv file. If the input data file uses a different separator between fields, then the CatReg functions `view` and `ascii` can be used to convert it to the comma-separated format. The `view` function can be used to read data from an external text file into an R data matrix. The `ascii` function then writes the contents of the data matrix to an external text (ascii) file. Both functions allow the user to specify the field separator using the option `sep="`*character string*`"`. Therefore, to convert a tab-separated file to a comma-separated file, proceed as in the following example.

```
> view(infile="data.tab", sep="\t") -> data
 Variables:
 mg/m3 SevLo
 View which variables? (all):
```

```
1:
> ascii(data, outfile="data.csv", sep=",")
```

The first command reads the external, tab-separated file into an R matrix named `data`.   The second command above writes the contents of `data` to an external comma-separated file called `data.csv`. This file will then be available for viewing using the view function or for analysis using `catreg`. The symbol "\t" represents the tab character. If fields are separated by a single space instead of by a tab character, specify `sep=" "` in the view function. The use of these functions is unnecessary if the input file is already comma delimited (i.e., uses a comma-separated format).

10.4.   Combining Severity Categories

Sometimes it is necessary to combine severity categories. For example, if some categories contain insufficient data, then adjacent severity categories may be combined to form broader categories. In some cases, the `catreg` or `partest` results may suggest that two or more of the severity parameters are not significantly different.

In these situations, the `join` function can be used to combine categories. This is a query-based function that allows specification of the lower and upper endpoints of an interval of categories to combine. The method for combining `SevLo` categories 1 and 2 for the data set `sim.csv` and writing the result to a new file `sim2.csv` is shown below.

```
> join(infile="sim.csv", outfile="sim2.csv")
Current observed categories are 0, 1, 2
Enter lower and upper scores for combined category
Lower:1
Upper:2
>>>> Output file is sim2.csv
```

If the `infile` and `outfile` names are the same, `join` displays a warning and quits. This prevents the data loss that would occur if the original file were overwritten by the modified file.

10.5. Recoding Variables

The CatReg `recode` function provides a way to recode specific variables. It works in a manner similar to the `join` function. When the `infile` and `outfile` arguments are specified, `recode` asks for the changes (see the example below).

```
> recode(infile="sim.csv", outfile="sim2.csv")
**** Recode specified variables ****
 Variables:
 mg/m3 SevLo
 Recode which variables? (none):
1: SevLo
2:
 Values of SevLo:
 0 1 2
 Enter replacement values:
1: 2 1 0
Name of new variable? (same):
```

In this example, the ordinal response was recoded from ascending order to descending order. Another use of the `recode` function is to recode $\{1,2,3,\dots,S\}$ as $\{0,1,2,\dots,S\text{-}1\}$.

10.6. Recoding Missing Values

Different systems often use different codes for missing values. The CatReg assumption is that a missing value is coded as a blank value. The CatReg `missval` function allows specification to CatReg of how missing values have been coded in the data and conversion of them to the form that CatReg expects. Suppose "-9999" is the code for missing data in `data.csv`; they are converted to blanks as shown below.

```
> missval(infile="data.csv", outfile="data2.csv")
**** Recode missing values ****
Enter the current code for missing data: -9999
```

After recoding the missing values, `missval` saves the modified data in the specified output file.

REFERENCES

Akaike, M. (1974) A new look at statistical model identification. IEEE Transactions on Automatic Control AU-19: 716-722.

Diggle, P. J.; Liang, K.-Y.; Zeger, S. L. (1994) Analysis of longitudinal data. New York, NY: Clarendon Press.

Huber, P. J. (1967) The behavior of maximum likelihood estimates under nonstandard conditions. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. Volume 1. Berkeley, CA: University of California Press.

Liang, K.-Y.; Zeger, S. L. (1986) Longitudinal data analysis using generalized linear models. Biometrika 73: 13-22.

Lopez, A.; Prior, M.; Yong, S.; Albassam, M.; Lillie, L. E. (1987) Biochemical and cytologic alterations in the respiratory tract of rats exposed for 4 hours to hydrogen sulfide. Fundam. Appl. Toxicol. 9: 753-762.

McCullagh, P.; Nelder, J. A. (1989) Generalized linear models. 2nd ed. London, United Kingdom: Chapman and Hall.

Morgan, B. J. T. (1992) Analysis of quantal response data. London, United Kingdom: Chapman & Hall.

Simpson, D. G.; Carroll, R. J.; Xie, M.; Guth, D. J. (1996a) Weighted logistic regression and robust analysis of diverse toxicology data. Commun. Stat. 25: 2615-2632.

Simpson, D. G.; Carroll, R. J.; Zhou, H.; Guth, D. J. (1996b) Interval censoring and marginal analysis in ordinal regression. J. Agric. Biol. Environ. Stat. 1: 354-376.

U.S. Environmental Protection Agency. (2000) CatReg software documentation. Research Triangle Park, NC: Office of Research and Development, National Center for Environmental Assessment; report no. EPA/600/R-98/053.

Venables, W. N.; Ripley, B. D. (1994) Modern applied statistics with S-Plus®. New York, NY: Springer-Verlag.

Weisberg, S. (1985) Applied linear regression. 2nd ed. New York, NY: Wiley.

APPENDIX A

DISTRIBUTION OF CONTINUOUS RESPONSE
DATA OVER SEVERITY LEVELS

Response data are sometimes measured on a continuous scale, with the mean and standard error reported. Section 3.2.3 contains a hypothetical example for lung weight in rats in which a mean and standard error were reported for each treatment group. The lung weight data were distributed over the severity levels shown for "Ref.id" = 2 in Table 3-6 for analysis by CatReg. The following description explains how that was done and provides an example calculation for reference. A technical explanation is provided at the end of the appendix.

The distribution of lung weights in healthy, unexposed rats is needed, either estimated from control animals in the experiment or "known" from other sources. It is assumed here that the distribution of lung weights is normally distributed with mean 1.0 g and standard deviation 0.05 g. (Note: The normal distribution is assumed simply for illustration. The same idea could be applied to other distributions). The user then determines weight intervals for severity levels to be used. For this purpose, it may be helpful to estimate first the highest weight that might be considered in the "normal" range for unexposed animals. The weight 1.15 g, which is three standard deviations above the mean, is an upper bound on virtually all lung weights in unexposed animals (i.e., a weight above 1.15 g is above the normal range of lung weights). The following correspondence was made between severity levels and lung weights for the example: Sev 0 (<1.15 g), Sev1 (1.15 to 1.50 g), and Sev2 (>1.50 g).

Suppose that the average lung weights and standard errors shown in Table A-1 were reported for treatment groups in "Ref.id" = 2, Table 3-6. (Note: If SE is the standard error from a treatment group of size n, the estimate of the standard deviation is $SE\sqrt{n}$.) There is a separate mean and standard error reported for each treatment group, denoted by $\overline{X}_i$, and SE_i, respectively for index i, with $i = 1,...,13$. The ith treatment group is assumed to be a sample from a normal distribution with unknown mean and standard deviation, denoted by μ_i and σ_i, respectively. The treatment groups are rather small ($n_i = 10$ for all i), so it was assumed that the standard deviation was the same for treatment groups with similar estimates of the standard

TABLE A-1. SUMMARY DATA FOR "REF.ID" = 2 IN TABLE 3-6

Index	Exp.	Group	mg/m3	Hours	Average Lung Weight ($\overline{X}$)	Standard Error
1	1	1	330	2	1.1	0.03
2	1	2	360	2	1.2	0.025
3	1	3	390	2	1.5	0.04
4	1	4	410	2	1.8	0.08
5	1	5	460	2	1.7	0.1
6	2	1	460	1	1.2	0.025
7	2	2	510	1	1.3	0.04
8	2	3	560	1	1.6	0.08
9	2	4	610	1	1.8	0.1
10	3	1	560	0.5	1.3	0.025
11	3	2	610	0.5	1.5	0.04
12	3	3	660	0.5	1.6	0.08
13	3	4	710	0.5	1.8	0.1

deviation (i.e., $SE_i \sqrt{n_i}$). The standard deviation was assumed equal for indices 1, 2, 6, and 10 (to be called Group A), 3, 7, and 11 (Group B), 4, 8, and 12 (Group C), and 5, 9, and 13 (Group D) in Table A-1. An estimate of the common standard deviation in Group A, σ_A, is calculated from SE_1, SE_2, SE_6, and SE_{10} as follows (estimation of the standard deviations for the other three groups is similar).

Let $\sigma_A{}^2$ be the common variance for Group A. The variance estimate from Sample 1 is $S_1{}^2 = n_1 SE_1{}^2 = 10\,(0.03)^2$. The estimate of $\sigma_A{}^2$, denoted by $S_A{}^2$ is the sum of the estimates of $\sigma_A{}^2$ from samples 1, 2, 6, and 10, weighted by their degrees of freedom (df) (9 for each sample). This gives $s_A{}^2 = 1/36\,[9(10)(0.03)^2 + 9(10)(0.025)^2 + 9(10)(0.025)^2 + 9(10)(0.025)^2] = 0.00694$. The proportion of Sample 1 with lung weights less than 1.15 g is estimated as follows. If X is a new observation for Sample 1, then

$$T = \left(X - \overline{X}_1\right) / \left(S_A^2\left(n_1 + 1\right) / n_1\right)^{0.5}$$

has a t-distribution with $n_A - 4$ df, where n_1 is the number of observations in Sample 1 (i.e., $n_1 = 10$), and n_A is the number of observations in Group A (i.e., $n_A = 40$). Pr ($X < 1.15$ g) is estimated by Pr ($T < (1.15 - 1.1)/[0.00694 \, (11/10)^{0.5}] = 0.716$, shown in Table A-2 under Sev0 for "Index" = 1. To estimate Pr ($1.15 \leq X < 1.5$), first estimate Pr ($X < 1.5$) as above, except with 1.15 g replaced by 1.5 g, and then subtract the estimate of Pr($X < 1.15$). Similarly, the relationship Pr($X \geq 1.5 = 1 - $ Pr($X < 1.5$) is used to estimate Pr ($X \geq 1.5$). The estimated proportions of each sample with lung weights in the intervals <1.15, 1.15 to 1.5, and $\geq$1.5 are displayed in Table A-2.

TABLE A-2. ESTIMATED PROPORTIONS IN SEVERITY LEVELS

Group	Index	$S\sqrt{\dfrac{n+1}{n}}$	$\overline{X}$	$\dfrac{1.15-\overline{X}}{S\sqrt{\dfrac{n+1}{n}}}$	df	Sev0 Pr $(X<1.15)$	$\dfrac{1.5-\overline{X}}{S\sqrt{\dfrac{n+1}{n}}}$	Sev1 Pr$(1.15<X<1.5)$	Sev2 Pr$(X>1.5)$
A	1	0.0874	1.1	0.572	36	0.715	4.577	0.285	0
A	2	0.0874	1.2	−0.572	36	0.285	3.432	0.714	0
B	3	0.1327	1.5	−2.638	27	0.007	0	0.493	0.500
C	4	0.2653	1.8	−2.450	27	0.011	−1.131	0.123	0.866
D	5	0.3317	1.7	−1.1658	27	0.127	−0.603	0.149	0.724
A	6	0.0874	1.2	−0.572	36	0.285	3.432	0.714	0.001
B	7	0.1327	1.3	−1.130	27	0.134	1.507	0.794	0.072
C	8	0.2653	1.6	−4.146	27	0	−0.377	0.354	0.645
D	9	0.3317	1.8	−1.960	27	0.030	−0.904	0.157	0.813
A	10	0.0874	1.3	−1.716	36	0.047	2.288	0.939	0.014
B	11	0.1327	1.5	−2.638	27	0.007	0	0.493	0.5
C	12	0.2653	1.6	−4.146	27	0	−0.377	0.354	0.645
D	13	0.3317	1.8	−1.960	27	0.030	−0.904	0.157	0.813

APPENDIX B

TECHNICAL DISCUSSION

B.1   LINK FUNCTIONS

Without loss of generality, link functions will be discussed in the context of Model 1, described by Eq. 4-1a in Section 4.1, specifically,

$$\Pr(Y \ge s|C,T) = H[\alpha_s + \beta_1 * f_1(C) + \beta_2 * f_2(T)].$$

CatReg currently supports three forms for H:

$$\text{Logistic} \quad H(x) = e^x / (1 + e^x),$$

$$\text{Normal} \quad H(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-z^2}{2}} dz \quad \textit{and}$$

$$\text{Gumbel} \quad H(x) = 1 - \exp(-e^x).$$

The inverse of H, which is denoted by L, is called the *link* function in the statistical literature. The link functions corresponding to the probability functions given above are

$$\text{Logit} \quad L(p) = \log[p / (1-p)],$$
$$\text{Probit} \quad L(p) = 100 \, p\text{th percentile of normal } (0,1), \text{ and}$$
$$\text{Cloglog} \quad L(p) = \log[-\log(1 - p)],$$

where p is any number between 0 and 1. The link function and probability function are inverse to each other in the sense that $H[L(p)] = p$ and $L[H(x)] = x$. Applying the link function to both sides of Model 1 gives

$$L[\Pr(Y \ge s \mid C,T)] = \alpha_s + \beta_1 * f_1(C) + \beta_2 * f_2(T),$$
$$s = 1,2,\ldots,S.$$

This expression shows that the link-transformed probability follows a linear model.

The use of a link function is essential here. Without it, the linear model becomes unbounded, and one is led to absurd estimates of probabilities for extreme values of C and T, namely negative probabilities or probabilities greater than 1. Moreover, link functions may be derived from a basic assumption that the ordinal severity score corresponds to exceeding an underlying toxic response threshold (see discussion below). Under this approach, the ordinal response score is called a quantal response because it is a quantization of the underlying response (see, e.g., Morgan, 1992). Any ordinal regression model of the type given in Model 1 may be interpreted as a quantal response model. When β_1 and β_2 are constant across severity categories data from one severity category adds information to the modeling of another category.

A toxicological interpretation of the link function follows from quantal response analysis. In particular, let Z denote a particular measure of health for a randomly selected subject given exposure to concentration C for duration T. Larger values of Z correspond to a healthier individual. Suppose that the health variable Z is distributed in the population as

$$\Pr(Z \leq x) = H\left(\frac{x - \mu}{\sigma}\right),$$

where x is a possible value for an individual's health, μ determines the median health for the given level of exposure, and σ is a measure of the population variation in health. In particular, if H is Gaussian, then the health levels in the population are distributed normally with mean μ and standard deviation σ. The larger μ is, the healthier the population is at the given level of exposure.

To model the situation where health deteriorates with increasing exposure, suppose that

$$\mu/\sigma = -\beta_{1*} f_1(C) - \beta_{2*} f_2(T).$$

Further, suppose that a toxic reaction of severity $Y \geq s$ occurs if the health measure Z is below a threshold $\sigma\alpha_s$, where σ is specific to the health measure. Then, under exposure (C, T), the probability of toxic severity of category s or higher is

$$\Pr(Y \geq s) = \Pr(Z \leq \sigma\alpha_s) = H[\alpha_s + \beta_{1*} f_1(C) + \beta_{2*} f_2(T)].$$

This is precisely the ordinal regression model given in Model 1. It is apparent that the link function is a reflection of the underlying distribution of Z, which cannot be measured directly, but its distribution, in particular the dependence on C and T, can be estimated indirectly from the toxicological response data.

B.2 INTERVAL CENSORING

Although censored responses do not provide as much information as fully scored responses, they do provide some information about the model. This information is used in the maximum likelihood estimation and the generalized maximum likelihood estimation described in Section B.3. In fitting the model by maximum likelihood, it is necessary to compute the probability of the observed response as a function of the model parameters. Table B-1 shows how these probabilities are computed in a three-category scoring system with interval censoring, using Model 1.

TABLE B-1. INTERVAL PROBABILITIES FOR MODEL 1 WITH THREE SEVERITY CATEGORIES

Interval	Probability
(0,0)	$1 - H[\alpha_1 + \beta_{1*} f_1(C) + \beta_{2*} f_2(T)]$
(0,1)	$1 - H[\alpha_2 + \beta_{1*} f_1(C) + \beta_{2*} f_2(T)]$
(1,1)	$H[\alpha_1 + \beta_{1*} f_1(C) + \beta_2 {}_* f_2(T)] - H[\alpha_2 + \beta_{1*} f_1(C) + \beta_2 {}_* f_2(T)]$
(1,2)	$H[\alpha_1 + \beta_{1*} f_1(C) + \beta_{2*} f_2(T)]$
(2,2)	$H[\alpha_2 + \beta_{1*} f_1(C) + \beta_2 {}_* f_2(T)]$

B.3 PARAMETER ESTIMATION

B.3.1 MAXIMUM LIKELIHOOD ESTIMATION

The *likelihood function* is defined to be the joint probability density of the data, viewed as a function of the parameters. In categorical regression, the response variables are discrete, so the likelihood may be interpreted as the probability that an investigation would result in the particular values that were observed. This probability depends on the unknown parameters. Maximum likelihood estimates the unknown parameters by the values that maximize the

likelihood of the observed data. It is often more convenient to work with the logarithm of the likelihood. For this purpose, it is common to define the *deviance* function:

$$\text{Deviance} = -2 * \log(\text{likelihood}).$$

The deviance is a nonnegative measure of model fit. Maximizing the likelihood is mathematically equivalent to minimizing the deviance. The factor 2 is included because it is the correct multiplier for certain likelihood-based, goodness-of-fit tests. Smaller deviances (larger likelihoods) correspond to a closer fit of the model to the data. A deviance of 0 would indicate a perfect fit, that is, a "saturated" model. Generally, a deviance of zero would indicate a model that is too complicated. The deviance value shown in CatReg summary statistics such as that in Table 5-7 is the "residual deviance" discussed in Section 7.

      If all data are independent, the likelihood function for interval-censored ordinal regression has a simple form. For $i = 1,2,..., n$, let Y_i denote the ordinal response, and let C_i and T_i denote the concentration and duration of exposure for the ith experimental subject, respectively. Y_i may be known only to lie in an interval. To account for this, let L_i and U_i denote the lower and upper endpoints of the known interval for Y_i, respectively. If it is known that $Y_i = k$, then $L_i = U_i = k$. For convenient reference, denote the probability of severity s or greater by

$$P_i(s) = \begin{cases} 1, & \text{if } s = 0; \\ H[\alpha_s + \beta_1 f_1(C) + \beta_2 f_2(T_i)], & \text{if } s = 1,\dots,S; \\ 0, & \text{if } s > S.. \end{cases}$$

Then the deviance for interval-censored ordinal regression is given by

$$\text{Deviance} = -2\sum_{i=1}^{N} \log[\Pr(L_i \leq Y_i \leq U_i | C_i, T_i)]$$
$$-2\sum_{i=1}^{N} \log[P_i(L_i) - P_i(U_i + 1)].$$

Parameter estimates are computed by iteratively minimizing the deviance. CatReg uses the R function `optim()` to perform this optimization.

B.3.2 GENERALIZED LIKELIHOOD ESTIMATION

Weighted ordinal regression analysis corresponds to a modified likelihood in which the probability associated with the ith observation is raised to a positive power w_i. This results in a modified likelihood with a weighted deviance:

$$\text{Deviance} = -2\sum_{i=1}^{N} w_i \log[P_i(L_i) - P_i(U_i + 1)].$$

If the weights do not correspond to incidences, then this likelihood corresponds to a nonstandard ordinal regression model. In this situation, it is more common to interpret the deviance as a generalized criterion and to assume that the usual ordinal regression model holds. Under this assumption, the generalized deviance still leads to consistent estimates of the parameters, but it does not correspond to the likelihood of the data. Instead, the estimator is defined by a generalized estimating equation, which provides the basis for computing valid large-sample confidence intervals and test statistics.

A further modification arises when the data are cluster sampled. The likelihood for cluster-sampled data does not have the simple form given above. Rather, it involves a product of multiple integrals of conditional likelihoods. Such likelihoods are computationally challenging and the results may be sensitive to the specification of the correlation structure. An alternative approach is to assume the ordinal regression model holds in a population-average sense. Consistent estimates then may be obtained quite generally, without making extensive distributional assumptions about the correlation structure. To achieve this, CatReg takes the expression derived above as a "working deviance" criterion. Minimizing it leads to consistent estimates under the population-average model. The main impact on the analysis compared to a standard likelihood analysis is the use of generalized estimating equation methods for making statistical inferences. In particular, rather than reporting the inverse information matrix as the estimated parameter covariance, the well-known sandwich formula is used.

Most applications of CatReg will involve cluster-sampled data and possible weighting of observations. As noted above, CatReg uses the weighted independence criterion as an estimating criterion, but computes confidence intervals and hypothesis tests without assuming the criterion is the likelihood. This approach has a long history in the statistical literature. Huber (1967) derived the large-sample theory of "maximum likelihood" estimators when the working likelihood is different from the actual likelihood of the data. In the literature on robust statistics,

this type of estimator is called an "M-estimator" because it generalizes maximum likelihood. Liang and Zeger (1986) extended the method to the analysis of correlated data, based on a "working" correlation structure, without assuming the working correlation structure was correct. This general approach is widely used in biometry, econometrics, and survey sampling.

B.4 CONFIDENCE LIMIT CALCULATIONS

CatReg uses the method of generalized estimating equations, which is well accepted in the literature (see Diggle et al., 1994), for the calculation of confidence limits for cluster-sampled data. The classical likelihood ratio inferences do not apply to cluster-correlated data because the likelihood ratio test assumes independent responses. The application of generalized likelihood ratio tests for correlated data is, however, an area of active research that may produce usable results in the future.

Confidence intervals and hypothesis tests about the parameters rely on a large-sample normal approximation to the joint distribution of the parameter estimates. The main steps in the derivation of this approximation are as follows. First, assuming the data are cluster sampled, write the generalized deviance as

$$GD = -2 \sum_{i=1}^{N} \sum_{j=1}^{n_i} w_{ij} \log[P_{ij}(L_{ij}) - P_{ij}(U_{ij}+1)],$$

where $-2\, w_{ij} \log[P_{ij}(L_{ij}) - P_{ij}(U_{ij}+1)]$ is the contribution of the jth individual from cluster i, $j = 1,...,n_i$, $i = 1,...,N$. Let B denote the vector of all model parameters. The GD estimate of B solves a generalized estimating equation

$$0 = \sum_{i=1}^{N} \sum_{j=1}^{n_i} \Psi_{ij},$$

where the summand Ψ_{ij} is given by

$$\Psi_{ij} = w_{ij} \frac{dP_{ij}(L_{ij}) - dP_{ij}(U_{ij}+1)}{P_{ij}(L_{ij}) - P_{ij}(U_{ij}+1)},$$

and $dP_{ij}(t)$ denotes the vector of derivatives of $P_{ij}(t)$, with respect to the components of B, evaluated at the estimated parameters. Expanding the estimating equation in a Taylor series leads to a large-sample normal approximation. The estimated parameter vector, $\hat{B}$, is approximately multivariate normal. The mean of the approximating normal distribution equals the true value of B, and the covariance matrix is given by the sandwich formula

$$\text{Est.Cov}(\hat{B}) = J^{-1}CJ^{-1},$$

where J is given by

$$J = -\sum_{i=1}^{N} \sum_{j=1}^{n_i} d\Psi_{ij},$$

and C is a covariance estimate for the total score, given by

$$C = \sum_{i=1}^{N} \left(\sum_{j=1}^{n_i} \Psi_{ij}\right)\left(\sum_{k=1}^{n_i} \Psi_{ik}\right)'$$

If the working likelihood were the actual likelihood of the data, then J and C would estimate the same matrices, and the usual inverse information, J^{-1}, would estimate the covariance of $\hat{B}$. Further details are given in Simpson et al. (1996a).

Standard errors of individual parameter estimates are obtained as the square roots of the diagonal elements of the estimated covariance. Confidence intervals and hypothesis tests derive from the normal approximation for $\hat{B}$.