residual analysis, model checking is based largely on fitting models that include extra terms (such as squared terms or interaction terms) whose significance would indicate shortcomings of the "target" model. Tests and confidence intervals for single coefficients may be carried out in a familiar way by comparing the coefficient estimate to its standard error. Tests of several coefficients are carried out by the drop-in-deviance chi-squared test. The drop in deviance is the sum of squared deviance residuals from the reduced model minus the sum of squared deviance residuals from the full model.

Donner Party

Given the likely inadequacy of the independence assumption and the inappropriateness of generalizing inferences from these data to any broader population, the analysis is necessarily informal. Nevertheless, it is useful to fit a logistic regression model for survival as a function of sex and age. Initial model fitting requires some examination into the linearity of the effect of age on the logit and of the interaction between sex and age. The coefficient of the sex indicator variable in the additive logistic regression model may be used to make a statement about the relative odds of survival of similarly aged males and females.

Birdkeeping and Lung Cancer

Initial tentative model fitting and model building are used to determine an adequate representation of the smoking variables for explaining lung cancer. Once years of smoking were included in the model, no further smoking variable was found significant when added. Then the variables associated with socioeconomic status and age were included, and interactions were explored. Finally, the effect of birdkeeping was inferred by adding the birdkeeping indicator variable into the model. A drop-in-deviance test established that birdkeeping had a significant effect, even after smoking was accounted for.

20.9 EXERCISES

Conceptual Exercises

1. **Donner Party.** One assumption underlying the correct use of logistic regression is that observations are independent of one another. (a) Is there some basis for thinking this assumption might be violated in the Donner Party data? (b) Write a logistic regression model for studying whether survival of a party member is associated with whether another adult member of the party had the same surname as that subject, after the effect of age has been accounted for. (c) Prepare a table of grouped responses that may be used for initial exploration of the question in part (b).

2. **Donner Party.** (a) Why should one be reluctant to draw conclusions about the ratio of male and female odds of survival for Donner Party members over 50? (b) What confounding variables might be present that could explain the significance of the sex indicator variable?

3. **Donner Party.** From the Donner Party data, the log odds of survival were estimated to be $1.6 - (0.078 \times age) + (1.60 \times fem)$, based on a binary response that takes the value 1 if an individual

survived and with *fem* being an indicator variable that takes the value 1 for females. (a) What would be the estimated equation for the log-odds of survival if the indicator variable for sex were 1 for males and 0 for females? (b) What would be the estimated equation for the log-odds of perishing if the binary response were 1 for a person who perished and 0 for a person who survived?

4. **Birdkeeping.** (a) Describe the retrospective sampling of the birdkeeping data. (b) What are the limitations of logistic regression from this type of sampling?

5. Since the term *regression* refers to the mean of a response variable as a function of one or more explanatory variables, why is it appropriate to use the word "regression" in the term *logistic regression* to describe a proportion or probability as a function of explanatory variables?

6. Give two reasons why the simple linear regression model is usually inappropriate for describing the regression of a binary response variable on a single explanatory variable.

7. How can logistic regression be used to test the hypothesis of equal odds in a 2 × 2 table of counts?

8. How can one obtain from the computer an estimate of the log-odds of $y = 1$ at a chosen configuration of explanatory variables, and its standard error?

Computational Exercises

9. **Donner Party.** It was estimated that the log odds of survival were $3.2 - (0.078 \times age)$ for females and $1.6 - (0.078 \times age)$ for males in the Donner Party. (a) What are the estimated probabilities of survival for men and women of ages 25 and 50? (b) What is the age at which the estimated probability of survival is 50% (i) for women and (ii) for men?

10. It was stated in Section 20.2 that if ω_A and ω_B are odds at A and B, respectively, and if the logistic regression model holds, then $\omega_A/\omega_B = \exp[\beta_1(A - B)]$. Show that this is true, using algebra.

11. **Space Shuttle.** The data in Display 20.14 are the launch temperatures (degrees Fahrenheit) and an indicator of O-ring failures for 24 space shuttle launches prior to the space shuttle *Challenger* disaster of January 28, 1986. (See the description in Section 4.1.1 for more background information.) (a) Fit the logistic regression of *Failure* (1 for failure) on *Temperature*. Report the estimated coefficients and their standard errors. (b) Test whether the coefficient of *Temperature* is 0, using Wald's test. Report a one-sided p-value (the alternative hypothesis is that the coefficient is negative; odds of failure decrease with increasing temperature). (c) Test whether the coefficient of *Temperature* is 0, using the drop-in-deviance test. (d) Give a 95% confidence interval for the coefficient of

DISPLAY 20.14 Launch temperature (degrees Fahrenheit) and incidence of O-ring failure for 24 space shuttle launches

Temperature	Failure	Temperature	Failure	Temperature	Failure
53	Yes	68	No	75	No
56	Yes	69	No	75	Yes
57	Yes	70	No	76	No
63	No	70	Yes	76	No
66	No	70	Yes	78	No
67	No	70	Yes	79	No
67	No	72	No	80	No
67	No	73	No	81	No

Temperature. (e) What is the estimated logit of failure probability at 31°F (the launch temperature on January 28, 1986)? What is the estimated probability of failure? (f) Why must the answer to part (e) be treated cautiously? (*Answer*: It is a prediction outside the range of the available explanatory variable values.)

12. Muscular Dystrophy. Duchenne Muscular Dystrophy (DMD) is a genetically transmitted disease, passed from a mother to her children. Boys with the disease usually die at a young age; but affected girls usually do not suffer symptoms, may unknowingly carry the disease, and may pass it to their offspring. It is believed that about 1 in 3,300 women are DMD carriers. A woman might suspect she is a carrier when a related male child develops the disease. Doctors must rely on some kind of test to detect the presence of the disease. The data in Display 20.15 are levels of two enzymes in the blood, creatine kinase (CK) and hemopexin (H), for 38 known DMD carriers and 82 women who are not carriers. (Data from D. F. Andrews and A. M. Herzberg, *Data*, New York: Springer-Verlag, 1985.) It is desired to use these data to obtain an equation for indicating whether a woman is a likely carrier.

DISPLAY 20.15 Values of creatine kinase (CK) and hemopexin (H) for 5 of the 82 controls (C = 0) and 5 of the 38 muscular dystrophy carriers (C = 1) in the muscular dystrophy data file

Controls (C=0)		Muscular dystrophy carriers (C=1)	
CK	H	CK	H
52	83.5	167	89
20	77	104	81
28	86.5	30	108
30	104	65	87
40	83	440	107

(a) Make a scatterplot of H versus log(CK); use one plotting symbol to represent the controls on the plot and another to represent the carriers. Does it appear from the plot that these enzymes might be useful predictors of whether a woman is a carrier? (b) Fit the logistic regression of carrier on CK and CK-squared. Does the CK-squared term significantly differ from 0? Next fit the logistic regression of carrier on log(CK) and $[\log(CK)]^2$. Does the squared term significantly differ from 0? Which scale (untransformed or log-transformed) seems more appropriate for CK? (c) Fit the logistic regression of carrier on log(CK) and H. Report the coefficients and standard errors. (d) Carry out a drop-in-deviance test for the hypothesis that neither log(CK) nor H are useful predictors of whether a woman is a carrier. (e) Typical values of CK and H are 80 and 85. Suppose that a suspected carrier has values of 300 and 100. What are the odds that she is a carrier relative to the odds that a woman with typical values (80 and 85) is a carrier?

13. Donner Party. Consider the Donner Party females (only) and the logistic regression model $\beta_0 + \beta_1 age$ for the logit of survival probability. If A^* represents the age at which the probability of survival is 0.5, then $\beta_0 + \beta_1 A^* = 0$ (since the logit is 0 when the probability is one-half). This implies that $\beta_0 = -\beta_1 A^*$. The hypothesis that $A^* = 30$ years may be tested by the drop-in-deviance test with the following reduced and full models for the logit:

$$\text{Reduced:} \quad -\beta_1 30 + \beta_1 age \quad = 0 + \beta_1(age - 30)$$
$$\text{Full:} \quad \beta_0 + \beta_1 age$$

To fit the reduced model, one must subtract 30 from the ages and drop the intercept term. The drop-in-deviance test statistic is computed in the usual way. Carry out the test that $A^* = 30$ for the Donner Party females. Report the two-sided p-value.

14. Donner Party. The estimate in Exercise 13 of A^* is $-b_0/b_1$. A 95% confidence interval for A^* can be obtained by finding numbers A_L^* and A_U^* below and above this estimate such that two-sided p-values for the tests $A^* = A_L^*$ and $A^* = A_U^*$ are both 0.05. This can be accomplished by trial and error. Choose several possible values for A_L^*, and follow the testing procedure described in the preceding exercise until a value is found for which the two-sided p-value is 0.05. Then repeat the process to find the upper bound A_U^*. Use this procedure to find a 95% confidence interval for the age at which the Donner Party female survival probability is 0.5.

Data Problems

15. Spotted Owl Habitat. A study examined the association between nesting locations of the Northern Spotted Owl and availability of mature forests. Wildlife biologists identified 30 nest sites. (Data from W. J. Ripple et al., "Old-growth and Mature Forests Near Spotted Owl Nests in Western Oregon," *Journal of Wildlife Management* 55: 316–18.) The researchers selected 30 other sites at random coordinates in the same forest. On the basis of aerial photographs, the percentage of mature forest (older than 80 years) was measured in various rings around each of the 60 sites, as shown in Display 20.16. (a) Apply two-sample t-tools to these data to see whether the percentage of mature forest is larger at nest sites than at random sites. (b) Construct a binary response variable that indicates whether a site is a nest site. Use logistic regression to investigate how large an area about the site has importance in distinguishing nest sites from random sites on the basis of mature forest. (Notice that this was a case–control study.) You may wish to transform the ring percentages to circle percentages.

DISPLAY 20.16 A subset of the Spotted Owl data: percentages of mature forest (>80 years) in successive rings around sites in western Oregon forests (5 of 30 random sites and 5 of 30 nest sites)

Randomly chosen sites: Outer radius of ring (km)							Spotted owl nest sites: Outer radius of ring (km)						
0.91	*1.18*	*1.40*	*1.60*	*1.77*	*2.41*	*3.38*	*0.91*	*1.18*	*1.40*	*1.60*	*1.77*	*2.41*	*3.38*
26.0	33.3	25.6	19.1	31.4	24.8	17.9	81.0	83.4	88.9	92.9	80.6	72.0	44.4
100.0	92.7	90.1	72.8	51.9	50.6	41.5	80.0	87.3	93.3	81.6	85.0	82.8	63.6
32.0	22.2	38.3	39.9	22.1	20.2	38.2	96.0	74.0	76.7	66.2	69.1	84.5	52.5
43.0	79.7	61.4	81.2	47.7	69.6	54.8	82.0	79.6	91.3	70.7	75.6	73.5	66.8
74.0	61.8	48.6	67.4	74.9	66.0	55.8	83.0	80.6	88.9	79.6	55.8	62.9	52.7

16. Bumpus Natural Selection Data. Hermon Bumpus analyzed various characteristics of some house sparrows that were found on the ground after a severe winter storm in 1898. Some of the sparrows survived and some perished. The data on male sparrows in Display 20.17 are survival status (1 = survived, 2 = perished), age (1 = adult, 2 = juvenile), the length from tip of beak to tip of tail (in mm), the alar extent (length from tip to tip of the extended wings, in mm), the weight in grams, the length of the head in mm, the length of the humerus (arm bone, in inches), the length of the femur (thigh bones, in inches), the length of the tibio-tarsus (leg bone, in inches), the breadth of

DISPLAY 20.17 A subset of data on 51 male sparrows that survived (SV = 1) and 36 that perished (SV = 2) during a severe winter storm: Age (*AG*) is 1 for adults, 2 for juveniles; *TL* is total length; *AE* is alar extent; *WT* is weight; *BH* is length of beak and head; *HL* is length of humerus; *FL* is length of femur; *TT* is length of tibio-tarsus; *SK* is width of skull; and *KL* is length of keel of sternum

SV	*AG*	*TL*	*AE*	*WT*	*BH*	*HL*	*FL*	*TT*	*SK*	*KL*
1	1	154	241	24.5	31.2	0.687	0.668	1.022	0.587	0.830
1	1	160	252	26.9	30.8	0.736	0.709	1.180	0.602	0.841
1	1	155	243	26.9	30.6	0.733	0.704	1.151	0.602	0.846
1	1	154	245	24.3	31.7	0.741	0.688	1.146	0.584	0.839
1	1	156	247	24.1	31.5	0.715	0.706	1.129	0.575	0.821
2	1	162	247	27.6	31.8	0.731	0.719	1.113	0.597	0.869
2	1	163	246	25.8	31.4	0.689	0.662	1.073	0.604	0.836
2	1	161	246	24.9	30.5	0.739	0.726	1.138	0.580	0.803
2	1	160	242	26.0	31.0	0.745	0.713	1.105	0.600	0.803
2	1	162	246	26.5	31.5	0.720	0.696	1.092	0.606	0.809

the skull in inches, and the length of the sternum in inches. (A subset of this data was discussed in Exercise 2.21.)

Analyze the data to see whether the probability of survival is associated with physical characteristics of the birds. This would be consistent, according to Bumpus, with the theory of natural selection: those that survived did so because of some superior physical traits. Realize that (i) the sampling is from a population of grounded sparrows, and (ii) physical measurements and survival are both random. (Thus, either could be treated as the response variable.)

17. Catholic Stance. The Catholic church has explicitly opposed authoritarian rule in some (but not all) Latin American countries. Although such action could be explained as a desire to counter repression or to increase the quality of life of its parishioners, A. J. Gill supplies evidence that the underlying reason may be competition from evangelical Protestant denominations. He compiled measures of: (1) *Repression* = Average civil rights score for the period of authoritarian rule until 1979; (2) *PQLI* (Physical Quality of Life Index in the mid-1970s) = Average of life expectancy at age 1, infant mortality, and literacy at age 15+; and (3) *Competition* = Percentage increase of competitive religious groups during the period 1900–1970. His goal was to determine which of these factors could be responsible for the church being pro-authoritarian in six countries and anti-authoritarian in six others. (Data from A. J. Gill, "Rendering unto Caesar? Religious Competition and Catholic Political Strategy in Latin America, 1962–1979," *American Journal of Political Science* 38(2) (1994): 403–25.)

To what extent do these measurements (see Display 20.18) distinguish between countries where the church took pro- and anti-authoritarian positions? Is it possible to determine the most influential variable(s) for distinguishing the groups?

18. Fatal Car Accidents Involving Tire Failure on Ford Explores. The Ford Explorer is a popular sports utility vehicle made in the United States and sold throughout the world. Early in its production, concern arose over a potential accident risk associated with tires of the prescribed size when the vehicle was carrying heavy loads, but the risk was thought to be acceptable if a low tire pressure was recommended. The problem was apparently exacerbated by a particular type of Firestone tire that was overly prone to tread separation, especially in warm temperatures. This type of tire was a common one used on Explorers in model years 1995 and later. By the end of 1999 more than

DISPLAY 20.18 Religious competition, quality of life, and repression in 12 predominantly Catholic Latin American countries

Catholic church stance	Country	PQLI	Repression	Competition (% increase)
Pro-authoritarian	Argentina	85	5.3	2.7
	Bolivia	39	4.3	4.1
	Guatemala	54	3.5	6.3
	Honduras	53	3.0	3.1
	Paraguay	75	5.2	2.1
	Uruguay	86	4.7	1.2
Anti-authoritarian	Brazil	66	4.8	12.0
	Chile	79	5.0	15.5
	Ecuador	69	3.7	2.9
	El Salvador	64	4.4	5.5
	Nicaragua	55	4.3	5.6
	Panama	79	5.7	4.4

DISPLAY 20.19 All 1995 and later model compact sports utility vehicles involved in fatal accidents in the United States between 1995 and 1999, excluding vehicles that were struck by other vehicles and those involved in alcohol-related accidents; first 5 of 2,321 rows

Type of sports utility vehicle (1 if Ford, 0 if not)	Vehicle age (years)	Number of passengers	Cause of fatal accident (1 if tires, 0 if not)
0	1	0	0
0	1	0	0
0	1	0	0
0	1	0	0
0	1	1	0

30 lawsuits had been filed over accidents that were thought to be associated with this problem. U.S. federal data on fatal car accidents were analyzed at that time, showing that the odds of a fatal accident being associated with tire failure were three times as great for Explorers as for other sports utility vehicles. Additional data from 1999 and additional variables may be used to further explore the odds ratio. Display 20.19 lists data on 1995 and later model compact sports utility vehicles involved in fatal accidents in the United States between 1995 and 1999, excluding those that were struck by another car and excluding accidents that, according to police reports, involved alcohol. It is of interest to see whether the odds that a fatal accident is tire-related depend on whether the vehicle is a Ford, after accounting for age of the car and number of passengers. Since the Ford tire problem may be due to the load carried, there is some interest in seeing whether the odds associated with a Ford depend on the number of passengers. (Suggestions: (i) Presumably, older tires are more likely to fail than newer ones. Although tire age is not available, vehicle age is an approximate substitute for it. Since many car owners replace their tires after the car is 3 to 5 years old, however, we may expect the odds of tire failure to increase with age up to some number of years, and then to perhaps decrease after that.

(ii) If there is an interactive effect of Ford and the number of passengers, it may be worthwhile to present an odds ratio separately for 0, 1, 2, 3, and 4 passengers.) The data are from the National Highway Traffic Safely Administration, Fatality Analysis Reporting System (http://www-fars.nhtsa.dot.gov/).

19. Missile Defenses. Following a successful test of an anti-ballistic missile (ABM) in July 1999, many prominent U.S. politicians called for the early deployment of a full ABM system. The scientific community was less enthusiastic about the efficacy of such a system. G. N. Lewis, T. A. Postol, and J. Pike ("Why National Missile Defense Won't Work," *Scientific American*, *281*(2): 36–41, August 1999) traced the history of ABM testing, reporting the results shown in Display 20.20. Do these data suggest any improvement in ABM test success probability over time?

DISPLAY 20.20 Date of ABM test, months after the first test, and result of the test; partial listing

Date	Months	Result
Mar-83	0	Failure
Jun-83	3	Failure
Dec-83	9	Failure
Jun-84	15	Success
...		
Jul-99	196	Success

20. Factors Associated with Self-Esteem. Reconsider the NLSY79 data in Exercise 12.23. The variable *Esteem1* is the respondent's degree of agreement in 2006 with the statement "I feel that I'm a person of worth, at least on equal basis with others," with values of 1, 2, 3, and 4 signifying strong agreement, agreement, disagreement, or strong disagreement, respectively. Construct a new binary response variable from this, which takes on the value 1 for strong agreement and 0 for agreement, disagreement, or strong disagreement. Explore the dependence of the probability that a person has positive self-esteem (as indicated by a response of strong agreement on *Esteem1*) on these explanatory variables: log annual income in 2005 (*Income2005*), intelligence as measured by the AFQT score in 1981 (*AFQT*), years of education (*Educ*), and gender (*Gender*).

Answers to Conceptual Exercises

1. (a) Yes. Members within the same family may have been likely to share the same fate; and if so, the binary responses for members within the same family would not be independent of one another. (It is difficult to investigate this possibility or to correct it for this small data set. Further discussion of this type of violation is provided in the next chapter.) (b) $\text{logit}(\pi) = \beta_0 + \beta_1 age + \beta_2 sur$, where $sur = 1$ if the individual had a surname that was shared by at least one other adult party member, and 0 if not. (c) Here is one possibility:

Age group	Number in group with shared surname	Proportion surviving	Number in age group without shared surname	Proportion surviving
15–25	13	0.625	8	0.375
26–45	10	0.800	7	0.000
>45	6	0.160	1	0.000

2. (a) There were no females over 50. Any comparisons for older people must be based on an assumption that the model extends to that region. Such an assumption cannot be verified with these

data. (b) If males and females had different tasks and if survival was associated with task, the tasks would be a confounding variable.

3. (a) $3.2 - 0.078\, age - 1.6\, male$. (b) Same equation but with all coefficients having opposite signs.

4. A sample of 49 individuals from all those with lung cancer was taken from a population. These were the "cases." Another sample of 98 individuals was taken from a similar population of individuals who did not have lung cancer. (b) Logistic regression may be used, with the 0–1 response representing lung cancer; but the intercept cannot be interpreted.

5. The mean of a binary response is a proportion (or a probability, if the populations are hypothetical).

6. (i) Proportions must fall between 0 and 1, and lines cross these boundaries. (ii) The variance is necessarily nonconstant.

7. Use a binary response to distinguish levels of the response category and an indicator variable to distinguish the levels of the explanatory category. The coefficient of the indicator variable is the log odds of one group minus the log odds of the other group. Carry out inference about that coefficient.

8. Subtract the specified value of each variable from all the variable values to create a new set of variables centered (zeroed) at the desired configuration. Run a logistic regression analysis on the new set, and record the "constant" coefficient and its standard error (see Section 10.2).