9.1 The diagnostic plots in Figures 9.2, 9.3, and 9.4 (on lecture slides) show three unusual observations in the Bankruptcy data. Fit a logistic regression model to the 63 observations without these three observations and compare your results with the results obtained in Section 9.3. Does the deletion of the three points cause a substantial change in the logistic regression results?

9.2 Examine the various logistic regression diagnostics obtained from fitting the logistic regression Y on X_1 and X_2 (Table 9.3 on lecture slides) and determine if the data contain unusual observations.

9.3 The *O-rings* in the booster rockets used in space launching play an important part in preventing rockets from exploding. Probabilities of O-ring failures are thought to be related to temperature. A detailed discussion of the background of the problem is found in *The Flight of the Space Shuttle Challenger* in Chatterjee and Simonoff (1995, pp. 33–35). Each flight has six O-rings that could be potentially damaged in a particular flight. The data from 23 flights are given in Table 1.11 and can also be found at the book's Website. For each flight we have the number of O-rings damaged and the temperature of the launch. Data available in the Canvas R codes folder

(a) Fit a logistic regression connecting the probability of an O-ring failure with temperature. Interpret the coefficients.

(b) The data for Flight 18 that was launched when the launch temperature was 75 degrees Fahrenheit was thought to be problematic, and was deleted. Fit a logistic regression to the reduced data set. Interpret the coefficients.

(c) From the fitted model, find the probability of an O-ring failure when the temperature at launch was 31 degrees Fahrenheit. This was the temperature forecast for the day of the launching of the fatal *Challenger* flight on January 20, 1986.

(d) Would you have advised the launching on that particular day?

Table 1.11 Number of O-rings Damaged and Temperature (Degrees Fahrenheit) at Time of Launch for 23 Flights of Space Shuttle *Challenger*

Flight	Damaged	Temperature	Flight	Damaged	Temperature
1	2	53	13	1	70
2	1	57	14	1	70
3	1	58	15	0	72
4	1	63	16	0	73
5	0	66	17	0	75
6	0	67	18	2	75
7	0	67	19	0	76
8	0	67	20	0	78
9	0	68	21	0	79
10	0	69	22	0	81
11	0	70	23	0	76
12	0	70			

9.4 Field-goal-kicking data for the entire American Football League (AFL) and National Football League (NFL) for the 1969 season are given in Table 9.15 and can also be found at the book's Website. Let $\pi(X)$ denote the probability of kicking a field goal from a distance of X yards.

(a) For each of the leagues, fit the model

Data available in the Canvas R codes folder

$$\pi(X) = \frac{e^{\beta_0+\beta_1 X+\beta_2 X^2}}{1+e^{\beta_0+\beta_1 X+\beta_2 X^2}}.$$

(b) Let Z be an indicator variable representing the league, that is,

$$Z = \begin{cases} 1, & \text{for the AFL,} \\ 0, & \text{for the NFL.} \end{cases}$$

Fit a single model combining the data from both leagues by extending the model to include the indicator variable Z; that is, fit

$$\pi(X, Z) = \frac{e^{\beta_0+\beta_1 X+\beta_2 X^2+\beta_3 Z}}{1+e^{\beta_0+\beta_1 X+\beta_2 X^2+\beta_3 Z}}.$$

(c) Does the quadratic term contribute significantly to the model?

(d) Are the probabilities of scoring field goals from a given distance the same for each league?

Table 9.15 Field-Goal-Kicking Performances of the American Football League (AFL) and National Football League (NFL) for the 1969 Season. The Variable Z is an Indicator Variable Representing League.

League	Distance	Success	Attempts	Z
NFL	14.5	68	77	0
NFL	24.5	74	95	0
NFL	34.5	61	113	0
NFL	44.5	38	138	0
NFL	52.0	2	38	0
AFL	14.5	62	67	1
AFL	24.5	49	70	1
AFL	34.5	43	79	1
AFL	44.5	25	82	1
AFL	52.0	7	24	1

Source: Morris and Rolph (1981, p. 200).

9.5 Using the data on diabetes analyzed in Tables 9.6 and 9.7 : (on lecture slides)

(a) Show that inclusion of the variable RW does not result in a substantial improvement in the classification rate from the multinomial logistic model using IR and SSPG.

(b) Fit an ordinal logistic model using RW, IR, and SSPG to explain CC. Show that there is no substantial improvement in fit, and the correct classification rate from a model using only IR and SSPG.