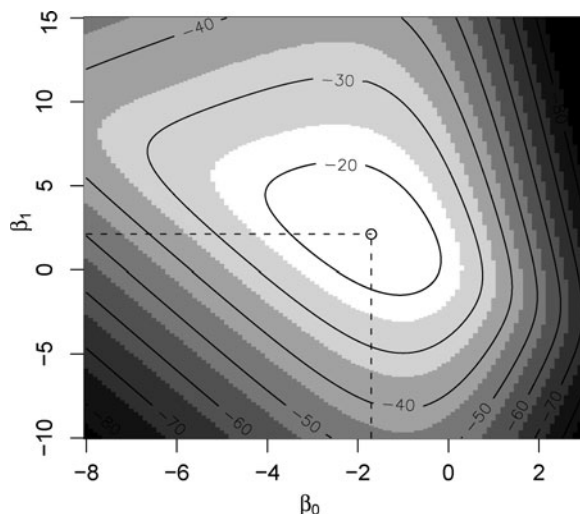


Fig. 5.8 Likelihood function for a two-parameter logistic model



Since both Y and x (the indicator of AIDS status) are observed, the only unknown quantities are the regression parameters β_0 and β_1 . These are generally estimated using an iterative maximization algorithm. Figure 5.8 presents a plot of the logarithm of this function for a range of values for P . Because the likelihood function depends on two unknown parameters, it has the form of a “surface” when plotted in three dimensions. The two-dimensional figure represents the contours of this surface as seen from above. The maximum value is indicated, and the corresponding maximum-likelihood estimates for β_0 and β_1 are -1.705 and 2.110 , respectively.

Because likelihoods are formed from the product of outcome probabilities for all individuals in a sample, the numerical value of a given likelihood depends on the sample size and is not particularly interpretable by itself. However, comparing likelihoods from nested models is a direct way to evaluate improvements in fit. This is the basis of the LR test.

Finally, we note that although the discussion here is limited to the binary outcome context, estimation methods for most of the regression models presented in this book are likelihood based. For example, least squares estimation and F -testing for comparing nested models in linear regression and analysis of variance models are examples of likelihood methods. Further, likelihood methods are fundamental to the family of GLMs discussed in Chap. 9.

5.7 Sample Size, Power, and Detectable Effects

Section 4.8 provides formulas for calculating sample size, power, and minimum detectable effects for the linear model. Analogous results hold for the logistic model. To compute the sample size that will provide power of γ in two-sided tests with

type-1 error of α to reject the null hypothesis $\beta_j = 0$ for the effect of a predictor X_j , accounting for the loss of precision due to multiple predictors, we can use

$$n = \frac{(z_{1-\alpha/2} + z_\gamma)^2}{(\beta_j^a \sigma_{x_j})^2 p(1-p)(1-\rho_j^2)}, \quad (5.16)$$

where β_j^a is the hypothesized value of β_j under the alternative, $z_{1-\alpha/2}$ and z_γ are the quantiles of the standard normal distribution corresponding to the specified type-1 error and power, σ_{x_j} is the standard deviation of X_j and ρ_j is its multiple correlation with the other covariates, and p is the marginal prevalence of the outcome (Hsieh et al. 1998; Hsieh and Lavori 2000). For problems with predetermined n , power is given by

$$\gamma = 1 - \Phi \left[z_{1-\alpha/2} - |\beta_j^a| \sigma_{x_j} \sqrt{np(1-p)(1-\rho_j^2)} \right] \quad (5.17)$$

and the minimum detectable effect (on the log-odds scale) by

$$\pm \beta_j^a = \frac{z_{1-\alpha/2} + z_\gamma}{\sigma_{x_j} \sqrt{np(1-p)(1-\rho_j^2)}}. \quad (5.18)$$

Some additional points:

- When X_j is binary with prevalence f_j , $\sigma_{x_j} = \sqrt{f_j(1-f_j)}$ in (5.16)–(5.18).
- When X_j is continuous with standard deviation σ_{x_j} , it is important to recognize that sample size, power, and minimum detectable effects do not depend in any real way on the units in which X_j is measured. This is most clearly seen in (6.26). Suppose X_j is usually measured in grams. Changing the unit to milligrams increases σ_{x_j} by a factor of 1,000, and shrinks β_j^a by the same factor. But of course the effect on the outcome of a 1-mg increase in the predictor is 1,000 times smaller than the effect of a 1-g increase. One way to avoid confusion is to consider the minimum detectable effect size for a one standard deviation change in X_j , which is often a reasonable-sized change to consider. That effect size is obtained by setting $\sigma_{x_j} = 1$ in (5.18).
- Sample size (5.16) and minimum detectable effect (5.18) calculations simplify considerably when we specify $\alpha = 0.05$ and $\gamma = 0.8$, β_j^a is the effect of a one standard deviation increase in continuous x_j , and we do not need to penalize for covariate adjustment. In that standard case,

$$n = \frac{7.849}{(\beta_j^a)^2 p(1-p)}. \quad (5.19)$$

For the minimum detectable effect, we have

$$\pm \beta_j^a = \frac{2.802}{\sqrt{np(1-p)}}. \quad (5.20)$$

For 90% power, substitute 10.51 for 7.849 and 3.242 for 2.802.

- Similarly, for a 2-arm clinical trial with equal allocation to arms, so that β_j^a is the log odds-ratio for treatment and $s_{x_j}^2 = 0.25$, we can calculate

$$n = 4 \times \frac{7.849}{(\beta_j^a)^2 p(1-p)}. \quad (5.21)$$

For the minimum detectable effect, we have

$$\pm \beta_j^a = 2 \times \frac{2.802}{\sqrt{np(1-p)}}. \quad (5.22)$$

- Power calculations using (5.17) simplify analogously, but still require a statistical calculator or computer package to evaluate the normal cumulative distribution function $\Phi(\cdot)$.
- As in calculations for the linear model, we need to use $|\beta_j^a|$ in (5.17) if $\beta_j^a < 0$. It follows that the negative of the value given by (5.18) is also a valid solution for the minimum detectable effect.
- These computations are valid for unmatched case-control studies, in which p , the sample prevalence of the outcome, is controlled by design. However, special methods are required for matched case-control studies.
- Sample size and power (but not minimum detectable effects) can be calculated using the `sampsi` command in Stata as well as many other statistical packages. Alternatively, (5.16)–(5.18) can easily be programmed in Stata, R, or Excel, or evaluated by hand if values of $z_{1-\alpha/2}$, z_γ , and $\Phi(\cdot)$ are available.
- The use of the factor $1 - \rho_j^2$ to account for covariate adjustment carries over from linear to logistic models. However, there is no analog to the reduction in residual variance that can result from including covariates in linear models, so that the adjustment to these calculations using $1 - \rho_j^2$ is less likely to be conservative.
- The `sampsi` command does not incorporate the factor $1 - \rho_j^2$ for covariate adjustment. An unadjusted estimate of n should be inflated by $1/(1 - \rho_j^2)$; similarly the unadjusted minimum detectable effect estimate should be inflated by $\sqrt{1/(1 - \rho_j^2)}$. To calculate power, use $n(1 - \rho_j^2)$ in place of n as an input.
- For logistic models with a continuous predictor, `sampsi` can be made to work by reversing the role of predictor and outcome, as we show in an example below.
- These calculations were derived in Chap. 4 from the Wald test of $\beta_j = 0$. Calculations based on the more reliable LR test (Self and Mauritsen 1992) have been implemented in the Egret statistical package.

- In Sect. 4.8, we showed how the standard error $SE(\hat{\beta}_j)$ plays a central role in sample size, power, and minimum detectable effect calculations for regression problems. $SE(\hat{\beta}_j)$ is a large-sample approximation for the logistic model, and more exact small-sample computations using the noncentral t -distribution do not carry over from the linear model. Simulations of power may be a more reliable guide when the calculated or available sample size is small.
- Equations (5.16)–(5.18) are based on the assumption that the conditional mean of the outcome does not vary strongly across observations, which would hold if X_j is a relatively weak predictor, or equivalently if $|\beta_j^a|$ is small. Methods based on simulation avoid this simplification and perform slightly better in some circumstances (Vittinghoff et al. 2009). However, errors from these sources are usually small compared to errors arising from uncertainty about the required inputs.
- The alternative calculations (4.22)–(4.24) presented in Sect. 4.8, which use an estimate $\tilde{SE}(\hat{\beta}_j)$ based on published results for an appropriately adjusted model using $\tilde{n}$ observations, carry over directly. There we showed that

$$n = \frac{(z_{1-\alpha/2} + z_\gamma)^2 \tilde{n} [\tilde{SE}(\hat{\beta}_j)]^2}{(\beta_j^a)^2}. \quad (5.23)$$

Similarly, power in a new sample of size n is given by

$$\gamma = 1 - \Phi \left[z_{1-\alpha/2} - |\beta_j^a| / [\sqrt{\tilde{n}/n} \tilde{SE}(\hat{\beta}_j)] \right]. \quad (5.24)$$

Finally, the minimum detectable effect in a new sample of size n can be obtained as

$$\pm \beta_j^a = (z_{1-\alpha/2} + z_\gamma) \sqrt{\tilde{n}/n} \tilde{SE}(\hat{\beta}_j). \quad (5.25)$$

In implementing these calculations, care must be taken to obtain the SE of the regression coefficient β_j , not the SE of the odds ratio e^{β_j} . Since results are usually available only for the odds ratio, this can be computed as $\tilde{SE}(\hat{\beta}_j) = \log(UL/LL)/3.92$, where UL and LL are the upper and lower 95% confidence bounds for the odds ratio. We must also ensure that X_j is measured on the same scale as in the published results.

To illustrate these methods, we first use the `sampsiz` command to estimate the sample size providing 80% power in two-sided tests with α of 5% for a clinical trial of a new technique hypothesized to reduce the incidence of an adverse postsurgical outcome from 15% to 5%. We specify the proportion with the outcome in each group, which are equivalent to the means of a continuous outcome. By omitting the `sd1` option, we signal that the outcome is binary, with SD determined under the statistical model as $\sqrt{p(1-p)}$. With equal allocation to treatment and control, the `r()` option, which specifies the ratio of the sizes of the groups being compared,

Table 5.33 Sample size calculation for randomized trial

```
. sampsi 0.05 0.15, power(0.8)

Estimated sample size for two-sample comparison of proportions

Test Ho: p1 = p2, where p1 is the proportion in population 1
           and p2 is the proportion in population 2

Assumptions:

      alpha =    0.0500   (two-sided)
      power =    0.8000
      p1     =    0.0500
      p2     =    0.1500
      n2/n1  =    1.00

Estimated required sample sizes:

      n1 =      160
      n2 =      160

. display log((0.05/0.95)/(0.15/0.85))
-1.2098379

. display (invnormal(0.975)+invnormal(.))^2/((-1.2098379)^2*0.25*0.075*
(1-0.075)) 309.17921
```

can be left at the default value of 1. In addition, we can safely assume that $\rho_j = 0$, so no adjustment for covariates is likely to be necessary in a randomized trial.

Table 5.33 shows the results. The `sampsi` command estimates that we need 160 participants per group. We also used (5.16) to estimate sample size. For that calculation, β_j^a , the hypothesized log-odds ratio for the effect of the new technique, is $\log(0.05/0.95)/(0.15/0.85) \approx -1.2098$. With equal allocation ($f = 0.5$) to treatment and control, $\sigma_x^2 = f(1 - f) = 0.25$, and the marginal prevalence p of the outcome $\approx 7.5\%$. This gives an overall sample size estimate of 309.

Now, suppose we would like to estimate the sample size that will provide 80% power in two-sided tests with α of 5% to detect an independent association of SBP with CHD, adjusting for age, smoking, BMI, cholesterol levels, and behavioral patterns, as suggested by the results in Table 5.10. From pilot data, we estimate that the prevalence of CHD in the new sample of high risk men will be 30%, that SBP will be approximately 5 mmHg higher among the men with CHD, that the within-group SD of SBP will be 15 mmHg, and finally that $\rho_j \approx 0.33$. To do this computation using the `sampsi` command, we reverse the role of the outcome and predictor, so f is now the prevalence of CHD. Pre-calculation of the local variable `ratio` is required because the `r()` option will not allow the fractional input 3/7.

Table 5.34 first shows the calculation using the `sampsi` command, with the adjustment using the variance inflation factor then carried out based on the unadjusted results. In addition, we computed the sample size using (5.16), relying on the fact that $\beta_j \sigma_x$, the log-odds per SD increase in SBP, is approximately equal to the standardized (in SDs) difference in mean SBP between the subgroups with and without CHD. The two sample size estimates are close.

Table 5.34 Sample size calculation for the effect of SBP on risk of CHD

```

. display 3/7

. 42857143

. sampsi 0 5, sd1(15) r(.42857143) power(0.8)

Estimated sample size for two-sample comparison of means

Test Ho: m1 = m2, where m1 is the mean in population 1
               and m2 is the mean in population 2

Assumptions:

      alpha =    0.0500   (two-sided)
      power =    0.8000
      m1 =           0
      m2 =           5
      sd1 =          15
      sd2 =          15
      n2/n1 =      0.43

Estimated required sample sizes:

      n1 =          236
      n2 =          102

. display (236+102)/(1-0.33^2) 379.30648

. display (invnormal(.975)+invnormal(0.8))^2/((5/15)^2*.3*.7*(1-.33^2))
377.48913

```

5.8 Summary

The logistic regression model extends frequency table techniques for investigating the association between a binary outcome and categorical predictor to include continuous predictors and allow simultaneous consideration of multiple (continuous and categorical) predictors.

Modeling techniques for logistic regression mirror those for linear regression, allowing many of the concepts and methods learned in Chap. 4 to be applied directly to studies involving binary outcomes. However, interpretation of logistic regression models is slightly more complex due to the model's nonlinear relationship between outcome risk and predictors. In particular, regression coefficients need to be transformed to be interpretable as odds ratios.

Although a powerful and useful tool, there are a number of situations where logistic regression is not the best method for analyzing binary outcome data. As we have seen in several examples, when attention is restricted to one or a few categorical predictors, regression techniques are not needed. In other situations, an alternative binary regression model linked to alternate measures of association such as relative risks or risk differences may be preferred. We refer readers to Chap. 6 for methods appropriate for regression analysis for event time outcomes. Although we have provided a brief illustration in Sect. 5.5.2 of how logistic regression can be used to investigate the effects of predictors on binary outcomes that are duration dependent,