zero observations, we need to use a model that explicitly allows for this. By the same token, if the zero observations are excluded from the sample, we need to take account of this omission. Both types of model are discussed in Section 11.6. The related problem of **sample selectivity**, in which certain observations are omitted from the sample in a nonrandom way, is dealt with in Section 11.7. Finally, in Section 11.8, we discuss **duration models**, which attempt to explain how much time elapses before some event occurs or some state changes.

11.2 Binary Response Models: Estimation

In a binary response model, the value of the dependent variable y_t can take on only two values, 0 and 1. Let P_t denote the probability that $y_t = 1$ conditional on the information set Ω_t, which consists of exogenous and predetermined variables. A binary response model serves to model this conditional probability. Since the values are 0 or 1, it is clear that P_t is also the expectation of y_t conditional on Ω_t:

$$P_t \equiv \Pr(y_t = 1 \,|\, \Omega_t) = \mathrm{E}(y_t \,|\, \Omega_t),$$

Thus a binary response model can also be thought of as modeling a conditional expectation.

For many types of dependent variable, we can use a regression model to model conditional expectations, but that is not a sensible thing to do in this case. Suppose that $\boldsymbol{X}_t$ denotes a row vector of dimension k of variables that belong to the information set Ω_t, almost always including a constant term or the equivalent. Then a linear regression model would specify $\mathrm{E}(y_t \,|\, \Omega_t)$ as $\boldsymbol{X}_t\boldsymbol{\beta}$. But such a model fails to impose the condition that $0 \leq \mathrm{E}(y_t \,|\, \Omega_t) \leq 1$, which must hold because $\mathrm{E}(y_t \,|\, \Omega_t)$ is a probability. Even if this condition happened to hold for all observations in a particular sample, it would always be easy to find values of $\boldsymbol{X}_t$ for which the estimated probability $\boldsymbol{X}_t\hat{\boldsymbol{\beta}}$ would be less than 0 or greater than 1.

Since it makes no sense to have estimated probabilities that are negative or greater than 1, simply regressing y_t on $\boldsymbol{X}_t$ is not an acceptable way to model the conditional expectation of a binary variable. However, as we will see in the next section, such a regression can provide some useful information, and it is therefore not a completely useless thing to do in the early stages of an empirical investigation.

Any reasonable binary response model must ensure that $\mathrm{E}(y_t \,|\, \Omega_t)$ lies in the 0-1 interval. In principle, there are many ways to do this. In practice, however, two very similar models are widely used. Both of these models ensure that $0 < P_t < 1$ by specifying that

$$P_t \equiv \mathrm{E}(y_t \,|\, \Omega_t) = F(\boldsymbol{X}_t\boldsymbol{\beta}). \tag{11.01}$$