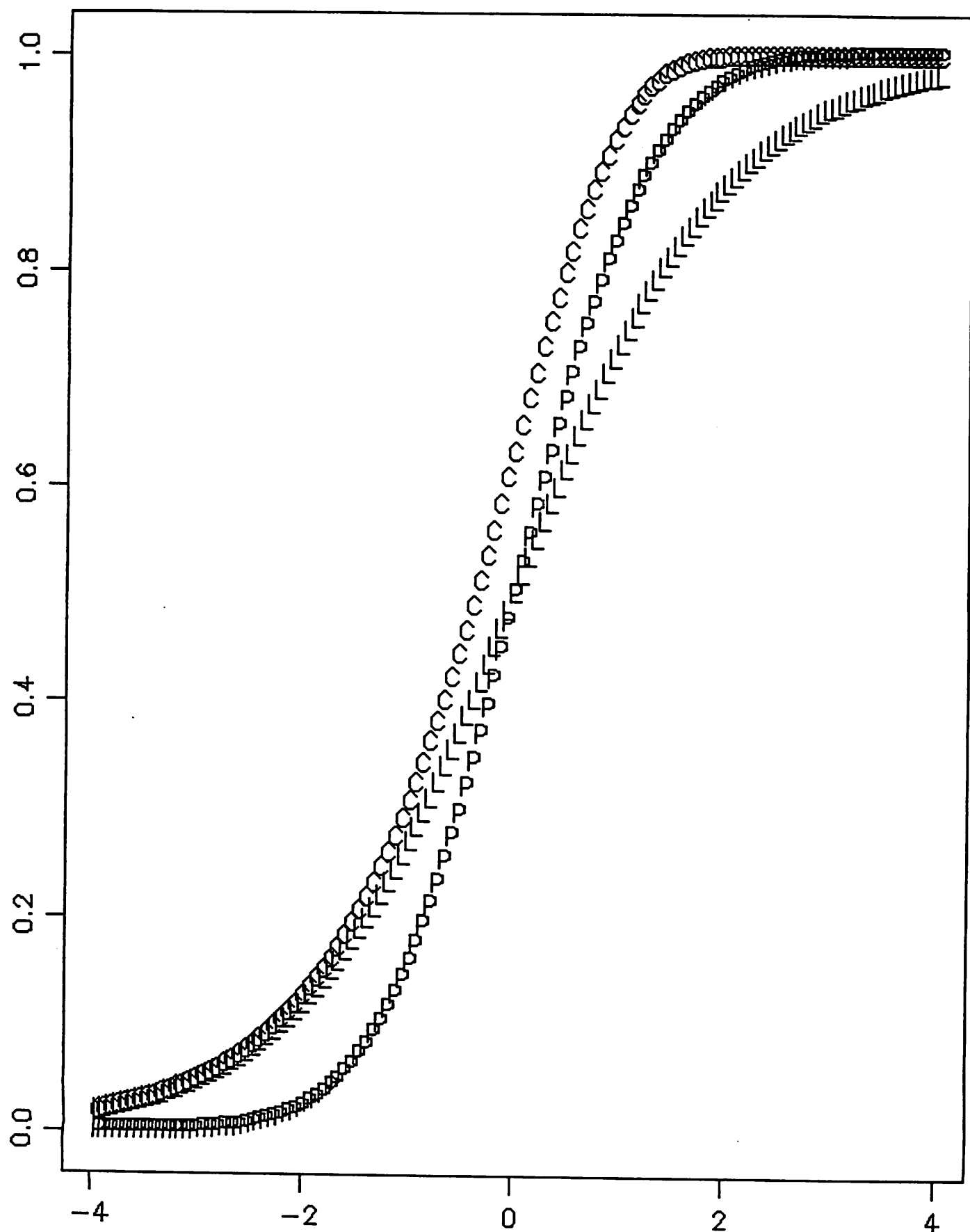


Binomial p models, comparing different link



linear predictor
Link functions: L=logit, P=probit, C=comp. log-log

linear combination (4.8) would be inconsistent with the laws of probability. A simple and effective way of avoiding this difficulty is to use a transformation $g(\pi)$ that maps the unit interval onto the whole real line $(-\infty, \infty)$. This remedy leads to instances of generalized linear models in which the systematic part is

$$g(\pi_i) = \eta_i = \sum_{j=1}^p x_{ij}\beta_j; \quad i = 1, \dots, n. \quad (4.9)$$

A wide choice of link functions $g(\pi)$ is available. Three functions commonly used in practice are

1. the logit or logistic function

$$g_1(\pi) = \log\{\pi/(1 - \pi)\};$$

2. the probit or inverse Normal function

$$g_2(\pi) = \Phi^{-1}(\pi);$$

3. the complementary log-log function

$$g_3(\pi) = \log\{-\log(1 - \pi)\}.$$

A fourth possibility, the log-log function

$$g_4(\pi) = -\log\{-\log(\pi)\},$$

which is the natural counterpart of the complementary log-log function, is seldom used because its behaviour is inappropriate for $\pi < \frac{1}{2}$, the region that is usually of interest. All four functions can be obtained as the inverses of well-known cumulative distribution functions having support on the entire real axis. The corresponding density functions are discussed in Exercises 4.22-4.23. The first two functions are symmetrical in the sense that

$$g_1(\pi) = -g_1(1 - \pi).$$

The latter two functions are not symmetrical in this sense, but are related via

$$g_3(\pi) = -g_4(1 - \pi).$$

All four functions are continuous and increasing on $(0, 1)$.

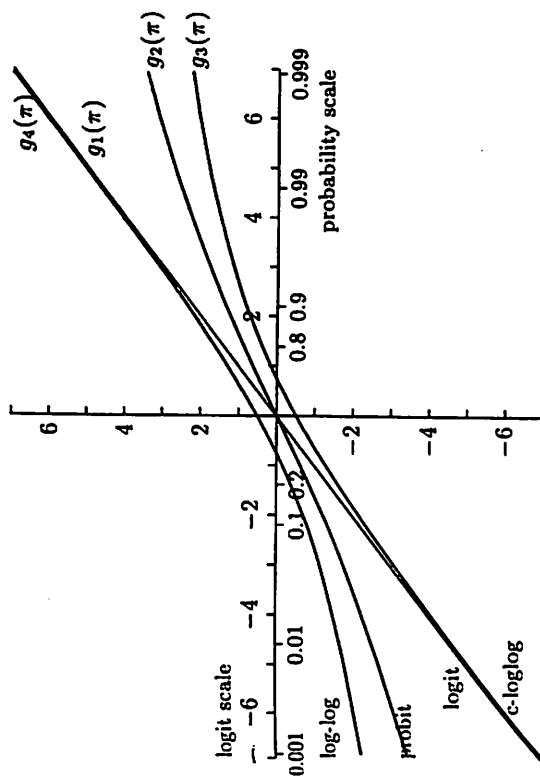


Fig. 4.1. A graphical comparison of three link functions with the logistic function: the 45° line is the logistic function.

Figure 4.1 compares the four functions. The logistic function is taken as the standard and $g_2(\pi)$, $g_3(\pi)$, $g_4(\pi)$ are plotted against $g_1(\pi)$ for values of π in the range 0.01 to 0.99.

The logistic and the probit function are almost linearly related over the interval $0.1 \leq \pi \leq 0.9$. For this reason, it is usually difficult to discriminate between these two functions on the grounds of goodness-of-fit; see, for example, Chambers and Cox (1967). For small values of π , the complementary log-log function is close to the logistic, both being close to $\log(\pi)$. As π approaches 1, the complementary log-log function approaches infinity much more slowly than either the logistic or the probit function. Similar comments apply to the log-log function as can be seen from Figure 4.1.

All asymptotic and approximate theory presented in this chapter applies regardless of the choice of link function. However, we shall be concerned mostly with the logistic function, not so much because of its simpler theoretical properties, but because of its simple interpretation as the logarithm of the odds ratio. Apart from this, the logistic function has one important advantage over all alternative transforms in that it is eminently suited for the analysis of data collected retrospectively.