

Questions

A researcher constructed a logit model for a binary dependent variable y_i and (s)he wants to use this model for forecasting. Assume that the logit model is the true model and that the used sample is large enough to ensure that the estimated parameters correspond to their true value.

(S)he constructs forecasted probabilities denoted by \hat{p}_i for m new values of y_i and wants to transform these values to 0/1 predictions using a cut-off value c . In this exercise we are going to derive which value of c the researcher has to choose to maximize the expected hit rate.

Consider the following pay-off matrix for the four situations that can occur:

observed	predicted	
	$\hat{y} = 0$	$\hat{y} = 1$
$y = 0$	1	0
$y = 1$	0	1

When the researcher predicts a one (zero) and the true value of y is zero (one) the pay-off is 0. When (s)he predicts a one or zero and the prediction is correct, the pay-off is 1.

- What is the expected pay-off in case the researcher predicts a one? (Hint. The true value y is a one with probability \hat{p}_i)
- What is the expected pay-off in case the researcher predicts a zero?
- Use the results in (a) and (b) to derive the optimal threshold to predict a one.