Calibration Curves:

Calibration curves are useful to quickly understand how well a model is classifying data. Calibration curves give a good measure of model performance over large amounts of data. A good calibration curve will be along the 45 degree line while the scale of both axes are 0 to 1. The x axis will be predicted probability, while the y axis will be empirical probability. The 45 degree line represents a 1 to 1 relationship between empirical and predicted probability. Points on the calibration curve are also commonly placed into bins, such as grouping probabilities 20-30% into one point placed at 25%.

Question 1:

TRUE/FALSE: If the event probability is 50%, we expect the model to correctly classify an observation as event 50% of the time

Answer: True, because the model will not be able to reach a better performance than the true probability.

Question 2:

TRUE/FALSE: When graphing predicted probability vs empirical probability, a well calibrated model will rise sharply and then flatten as x increases.

Answer: False, that is the behavior of an ideal ROC curve. A well calibrated model should lie along the 45 degree line, representing a 1 to 1 relationship.