averaged hinge loss in test set (IQ)