



***“If model selection and true error estimates are to be computed simultaneously, the data should be divided into three disjoint sets”***

***[Brian D. Ripley, 1996]***

**1. Training set (a.k.a. pseudo-training)**

**2. Validation set (a.k.a. pseudo-test)**

**3. Test set (*unseen!*)**

**“If model selection and true error estimates are to be computed simultaneously, the data should be divided into three disjoint sets”**

**[Brian D. Ripley, 1996]**

1. Training set (a.k.a. pseudo-training)
2. Validation set (a.k.a. pseudo-test)
3. Test set (*unseen!*)

# K-Fold Cross Validation

cross validation **simulates multiple train-test-splits** on the training data