- The bootstrap increases the variance that can occur in each fold (desirable)
- Example: a classification problem
  - ‣ Assume C classes, N observations, N_C observations for each class
  - ‣ The a *priori* probability of choosing an example from class C is N_C/C
  - ‣ Without replacement, once class is chosen, probability has changed to $(N_C-1)/C$
- Sampling with replacement preserves a priori probabilities of random selection

Comparison CV and Bootstrapping

Comparison CV and Bootstrapping

- The bootstrap increases the variance that can occur in each fold (desirable)
- Example: a classification problem
  - Assume C classes, N observations, N_C observations for each class
  - The a *priori* probability of choosing an example from class C is N_C/C
  - Without replacement, once class is chosen, probability has changed to $(N_C-1)/C$
- Sampling with replacement preserves a priori probabilities of random selection

Cross Validation and Panel Data