Resampling Methods as Applied to Significance Tests: A Case Study

Veronica Powell, Manager, Biostatistics
Carrie Caswell, Statistician

Resampling methods are an effective tool for calculating inferential statistics from empirical data. We performed the analysis for a reader training study in which inter-reader agreement was evaluated with Fleiss's Kappa statistic. Subjects were injected with a contrast agent, and physicians rated brain scans based on presence or absence of a disease. Physicians then underwent a training process and read the scans again. We analyzed the agreement both before and after training, and obtained inference statistics on the change.

Data: 96 subjects, 37 readers; each reader gave a positive or negative reading on each subject.

KAPPA: $\kappa = P_o - P_e / 1 - P_e$, a measure of the probability of inter-reader agreement corrected for random chance.
P_o=observed probability of agreement
P_e=probability of agreement by random chance

POINT ESTIMATE: $(\kappa_{After} - \kappa_{Before})$

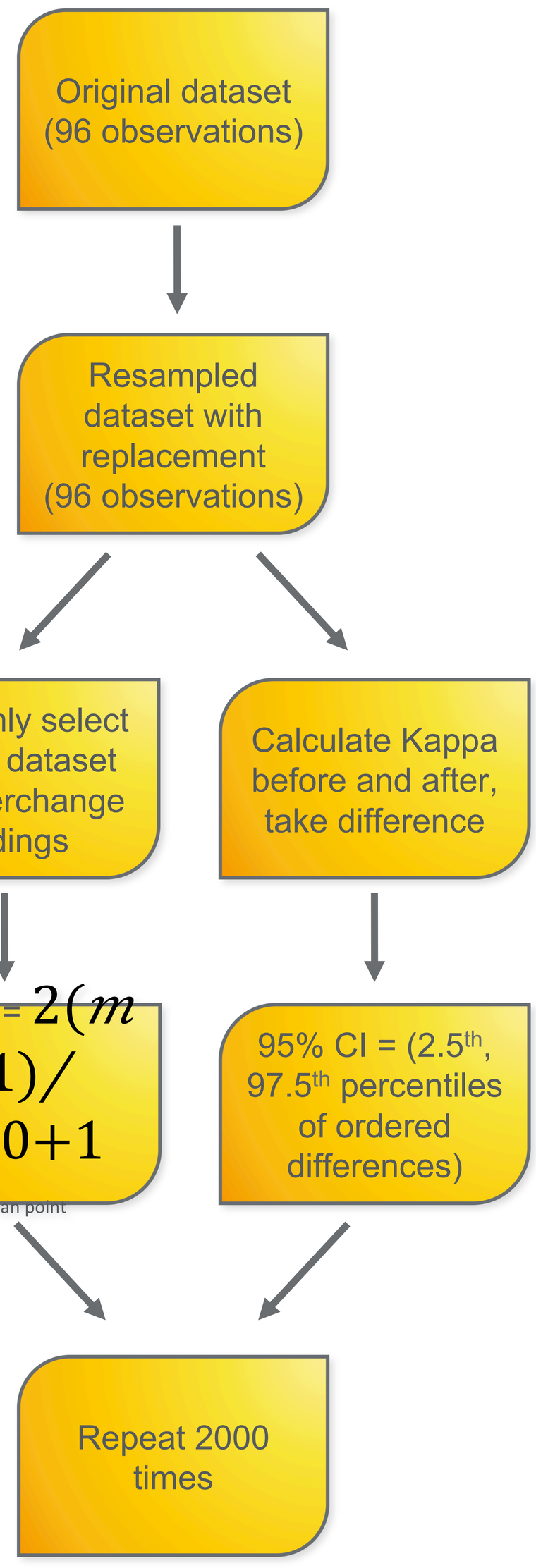
- How to obtain inference statistics on this?
- Resample!
- We need:
  - Confidence interval for change in Kappas (Bootstrapping method)
  - P-value for change in Kappas (Monte Carlo resampling)

CHALLENGES:

- ***Run time: 11 hours***
  - Program was run on a Unix system.
  - Minor mistakes take days to debug or even detect.
  - **Solution:** Instead of one program consisting of 5 macro calls, split into 5 programs which can be run separately.
  - Each takes 1-2 hours to run, and errors can be detected and debugged in subsequent programs.
  - **Possible Solutions:** More efficient use of macros, techniques such as *proc append* in place of *set* statements.
- ***P-value requires a dynamically allocated seed***
  - Randomly selecting 96 observations with replacement 2000 times can be successfully completed with a static seed.
  - We accomplished this by selecting all (2000*96)=192,000 observations in one procedure, then later separating into smaller datasets.

```
proc surveyselect data=orig out=resamp seed=1234 n=96 reps=2000
method=urs;
run;
```

  - The first draft of this program used seed '1234' for the second random selection (50% without replacement of the resampled data).
  - The same 48 observations were selected every time!
  - **Solution:** Allocate a string of 2000 random seeds, one for each replicate.
- ***Validation Code and Production Code cannot always match p-values to more than 3 decimals***
  - Production uses a pre-defined SAS macro (%magree) for all Kappa calculations; validation calculates by hand.
- ***These challenges and solutions can be applied to any resampling methods in Unix-based SAS***



THEOREM CLINICAL RESEARCH™