Sample Size Designs to Assess Controls

B. Ricky Rambharat, PhD, PStat®
Lead Statistician

Office of the Comptroller of the Currency
U.S. Department of the Treasury
Washington, DC

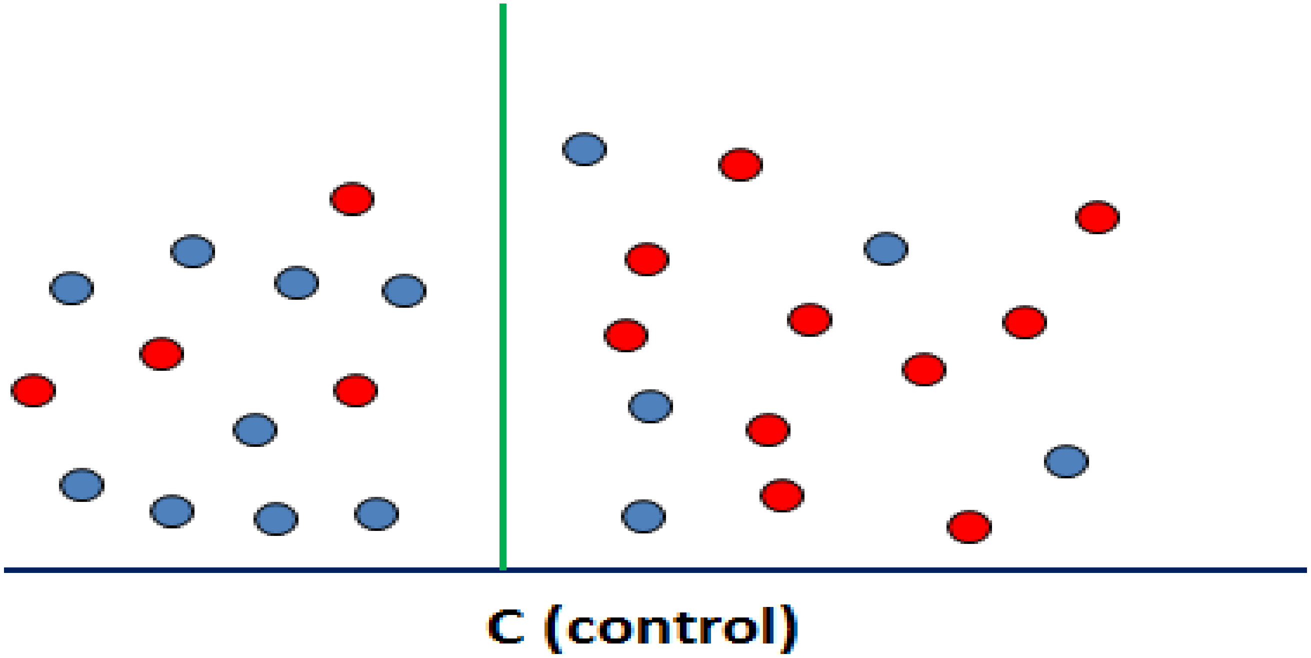
FCSM Research Conference
Session B-2

Nov. 4th, 2013

Disclaimer

The views expressed in this presentation are solely those of the author. This talk neither represents the opinions of the U.S. Department of the Treasury nor the Office of the Comptroller of the Currency (OCC). All errors are author's responsibility. As this talk is based on a paper in **draft status**, please do not reference without conferring with the author.

Motivation: How to Set Control?



Examples of Controls

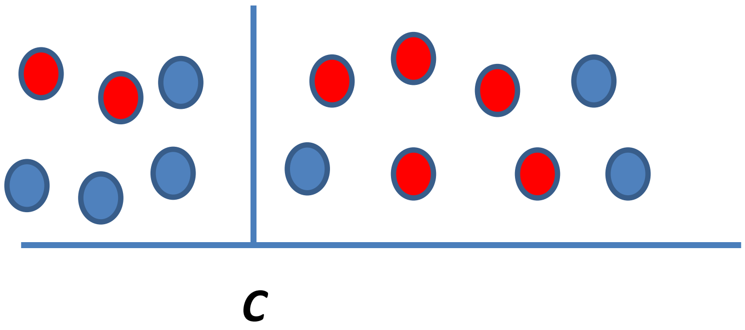
- Currency Transaction Report (CTR)
  - Flag if > US $10,000.00
- Credit score cutoffs
  - Accept / Deny applicant
- College applications and entrance exams
- Job applications and interview process
- **Challenge:** control-setting and risk

Type I vs. Type II Error Control

- Type I error: False positives
  - Compliance costs
  - Manpower
  - Resource constraints
  - Job offer to weak candidates
  - Accept weak students

- Type II error: False negatives
  - Criminal activity
  - Unfair consumer practices

Sample Design & Analysis



- *Goal:* Assess control *C*
- *Definition:* What is an "error"?
- *Q:* What sample size *n* to gauge Type II error?
- *Assumption:* "Rare error" view?
- *Result:* estimate error <u>rate</u> below *C*
- *Follow-on:* How to do deep-dives? 2-stage?

Sampling Models

- ***Hyper-Geometric*** *distribution*
  - Sampling <u>without</u> replacement model
- ***Normal*** *distribution* (approx.)
  - Cochran (1977)
  - Accommodate rare errors?
- ***Binomial*** *distribution* (approx.)
  - *n, p*
  - Any $0<p<1$; numerical solution
- ***Poisson*** *distribution* (approx.)
  - Related to Binomial

Inference I: Total Errors

- Estimation of population error rate: *p*
  - Hypothesis-testing problems & deep-dives
  - “Ballpark” notion of error incidence
  - Useful for regression modeling?
- Estimation of **<u>total</u> number of errors**: ***M***
  - Relevance to key policy decisions
  - A different view of risk
  - Informative for deep-dive studies?

Inference II: Total Errors

- Sampling models for *M* (discrete data)
  - Hyper-Geometric distribution
  - Binomial (alternative)
  - Normal (OK…could be inaccurate)

- Computations
  - Prob. of exactly ***k*** errors in a sample of size ***n***
  - Prob. of at least ***k*** errors in a sample of size ***n***

Bayesian Estimation

- **Goal:** Posterior distribution of M
  - How are total errors distributed?
  - Variability measures
  - Risk evaluations
  - Relevance for deep-dive analysis
  - Markov chain Monte Carlo

- **Alternative:** Posterior distribution of p
  - Beta-Binomial framework

Prior Distributions

- **Discrete Uniform**
  - Mass on {0,1,2,…,N}
  - Vague / Non-informative (reflects ignorance)
- **Poisson**
  - Mass on {0,1,2,3,…}
  - Leverage data
  - Empirical Bayes
- **Plan:** use *mixture* to construct prior, *p(M)*.
  - Balance “ignorance” and data.

Likelihood Function

- Quantities of interest
  - *Pr*(exactly *k* errors)
  - *Pr*(at least *k* errors)

- ***Hyper-Geometric (HG) likelihood***
  - Limited for large *N* (population size)
  - Workaround: Compute log-probabilities

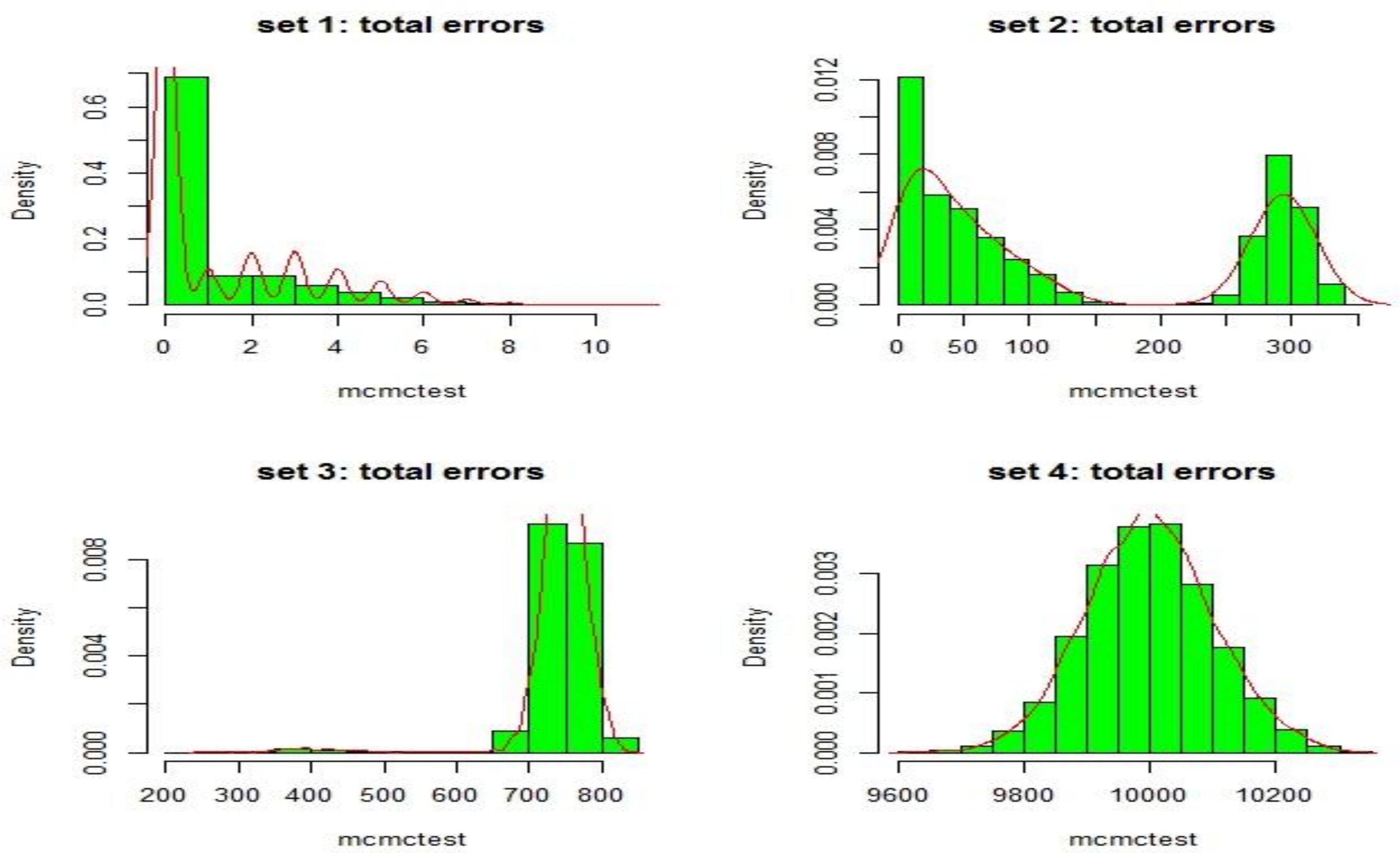
Proposal Distribution

- Poisson proposal distribution
- Choose $\boldsymbol{\lambda}$: another mixture?
- Weight $\boldsymbol{\lambda}$ (“ignorance” vs. “MCMC learning”)
- MCMC mixing properties
- Convergence of MCMC chain
- Satisfaction of ***detailed balance*** condition

Numerical Experiments

INPUTS	Set 1	Set 2	Set 3	Set 4
k (# of errors)	0	1	1	1
n (sample size)	100	100	40	100
N (pop. size)	30,000	30,000	30,000	1,000,000
Weights: Prior {DU; Pois}	{0.999; 0.001}	{0.01; 0.99}	{0.01; 0.99}	{0; 1}
Weights: Prop. {Fix; Dynamic}	{0.90; 0.10}	{0.01; 0.99}	{0.001; 0.999}	{0; 1}

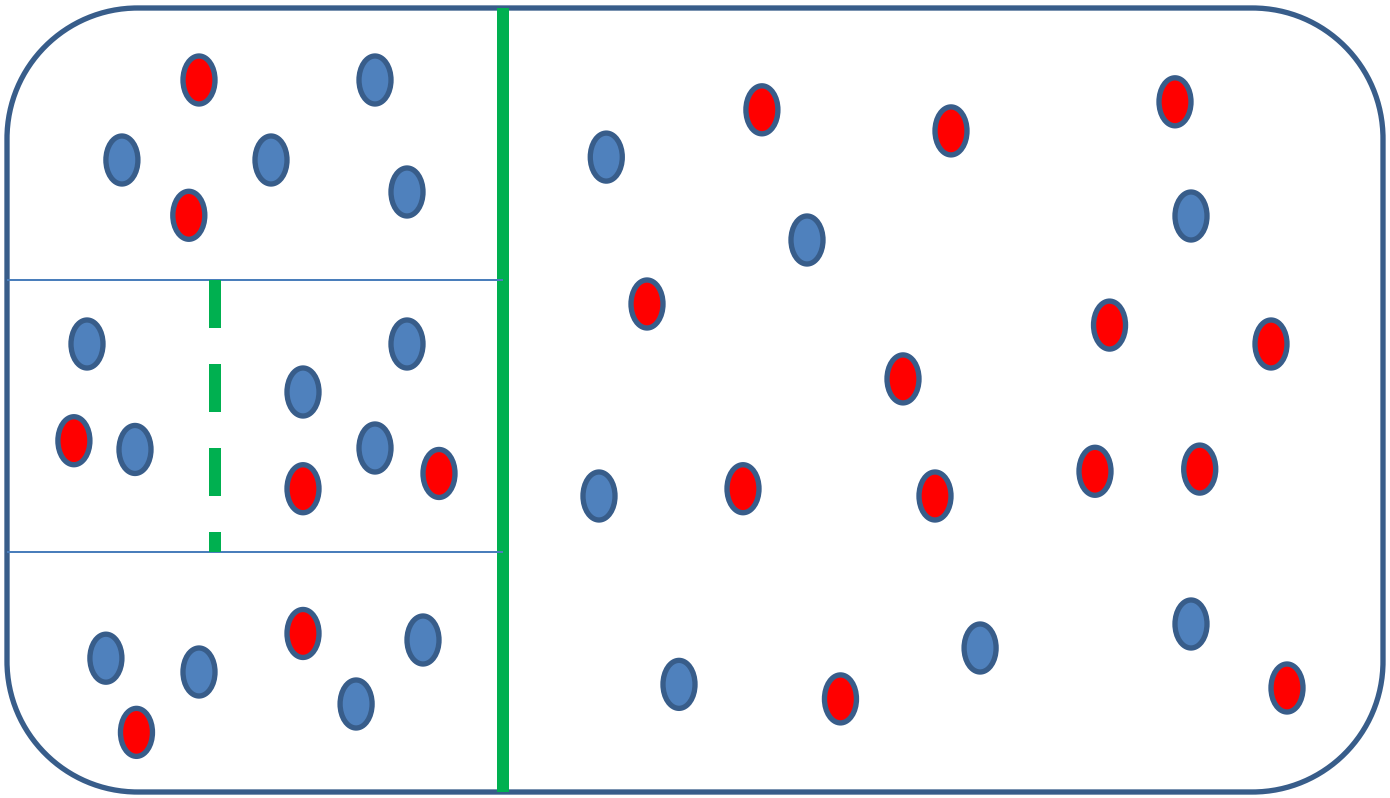
Posterior Results: Univariate Errors



Deep-Dive Sample Design

- Leverage stage 1 results for stage 2
- Posterior simulations and deep-dive
- Probabilistic view of risk
  - Easier for uni-modal posteriors
- Practical constraints vs. posterior outcome
  - E.g., bi-modal distributions
- Relevance for sample size $n = n(\text{posterior})$

Multinomial Experiment (Schematic)



Sample Size & Control-Setting

- Multinomial probabilities
  - Chi-squared tests of significance
- Deep-dive and sample size by segment
- Estimate total errors in risky sub-populations
  - Straightforward mapping to error rate
- Relevance to control settings

Multiple Levels of Errors

- Attribute sampling: binary error incidence
- E.g., degree of "injury" to consumer
- Posterior analysis
  - Bivariate HG likelihood
  - Bivariate Poisson proposal distribution
  - Co-dependence between errors?
- General modeling
  - E.g., Ordinal, Multinomial logit

Preliminary Conclusions

- Total number of errors vs. Error rate
- Posterior analysis of total errors
- Subjective / Data influence
- Sample design and posterior results
- Link control settings to risky sub-populations

Future Work

- Framework for multiple levels of errors
- MCMC estimation / posterior analysis
- Sample size and posterior results
- Logic for control settings
- Computational tool

END: THANK YOU