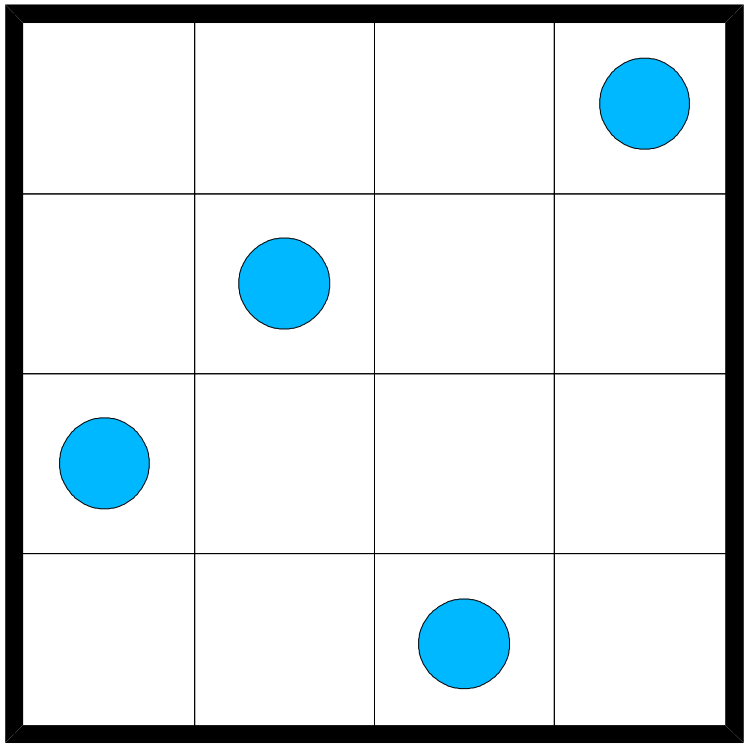
Acceptance/Rejection Sampling from Restricted Permutations

Mark Huber
NSF Postdoc
Visiting--Department of Statistics
Stanford University
mhuber@orie.cornell.edu
www-stat.stanford.edu/~mhuber

Rook Placements



1 2 3 4

4 2 1 3

The Problems

- Generation
  - Uniformly from set of restricted permutations
  - Previous: Markov chain
- Counting
  - Permanent of 0-1 matrix $A=(a_{ij})$

$$per(A)=\sum_{\sigma} a_{i\,\sigma(i)}$$

  - Randomized algorithms

Applications

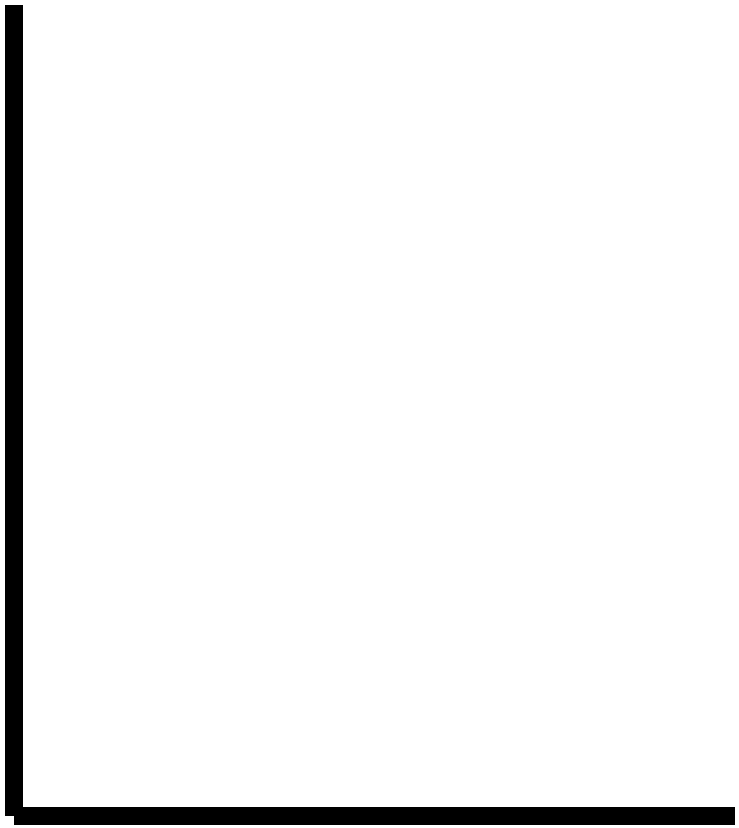
- Permanent of 0-1 matrix
  - #P complete
  - Ryser Formula

$$O\left(n2^n\right)$$

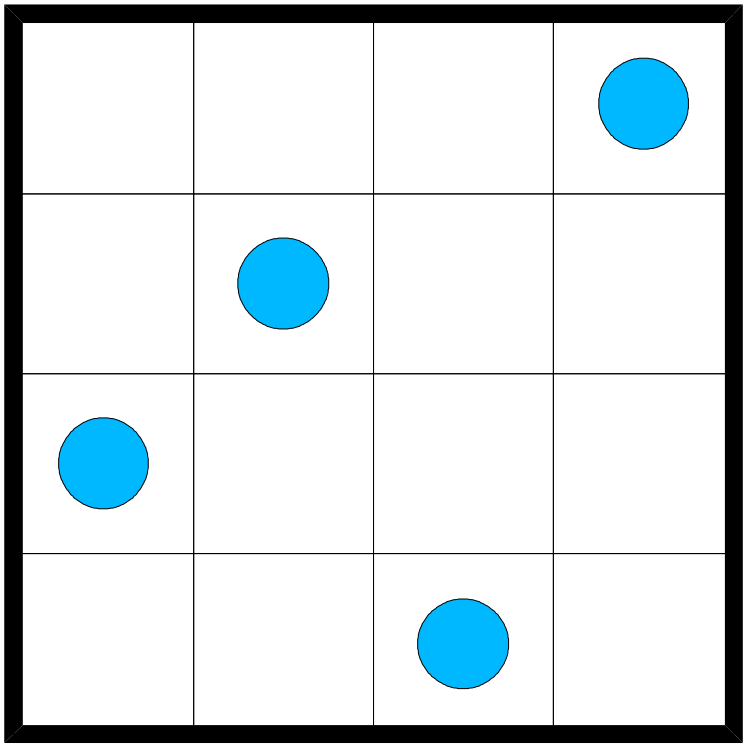
- Latin Rectangles
- Doubly Truncated Data
  - Quasar Luminosity Evolution

Quasar Luminosity Evolution

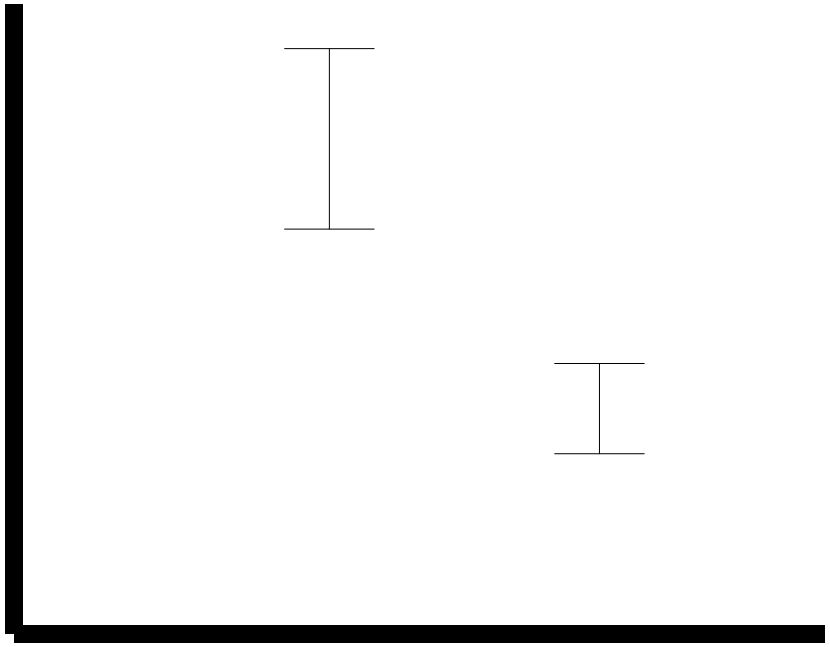
Kendall's Tau



Number of inversions

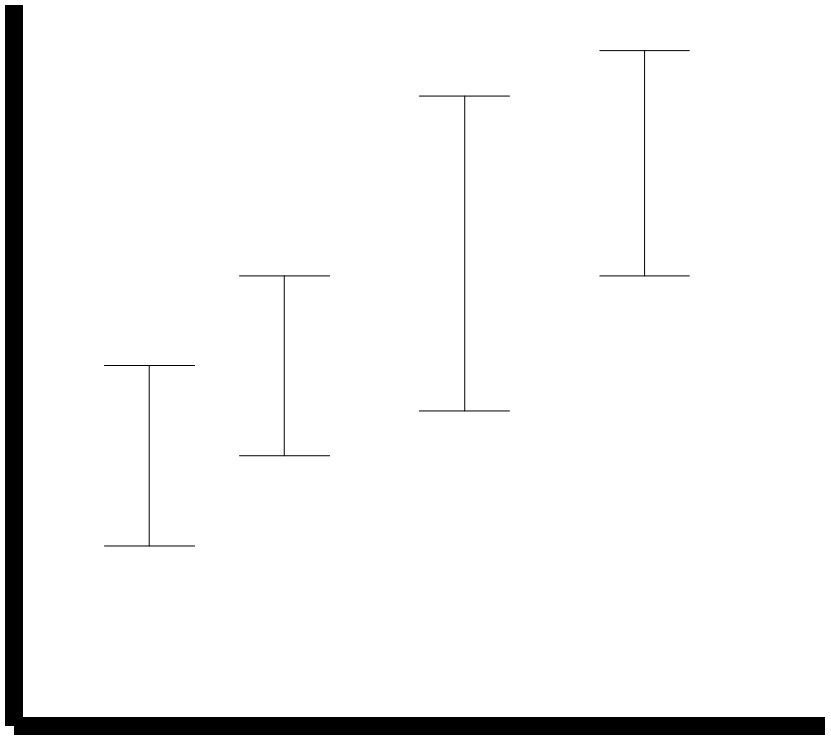


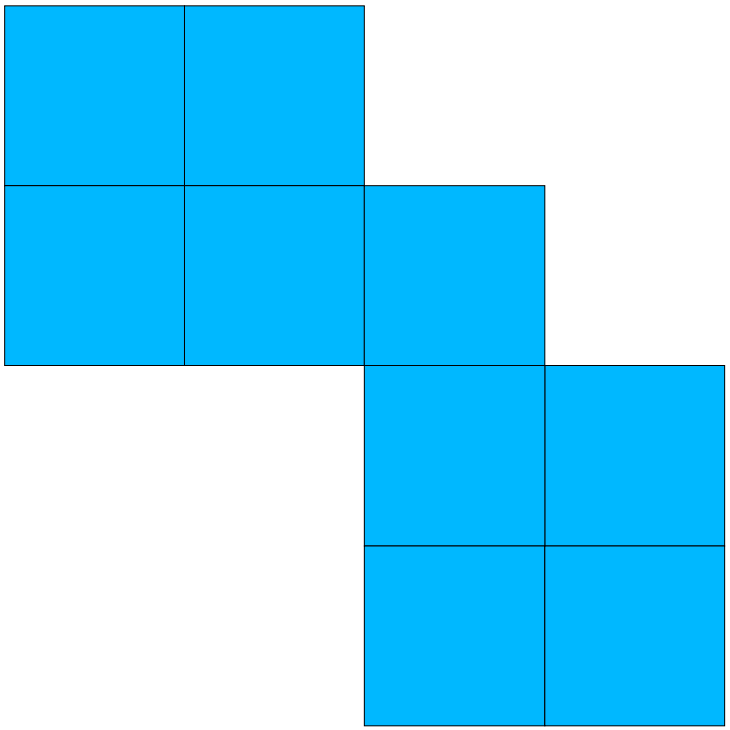
Doubly Truncated Data



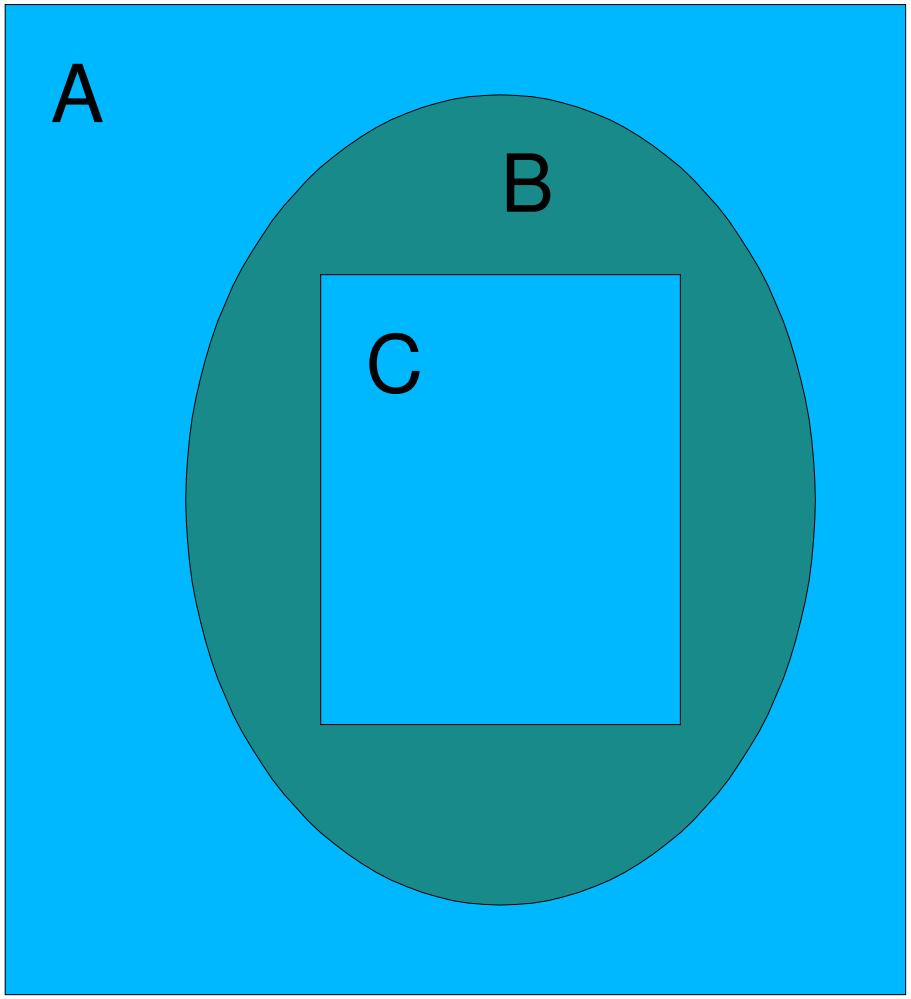
$$\tilde{\tau} = \sum sign[(i-j)(\sigma(i)-\sigma(j))]$$

Interval Permuations





Acceptance/Rejection



Easy to sample from A

Prob. of acceptance

$$\frac{|B|}{|A|}$$

Running Time

$$E[T] = \frac{|A|}{|B|} \leq \frac{|A|}{|C|}$$

Two Approaches

Minc's Inequality

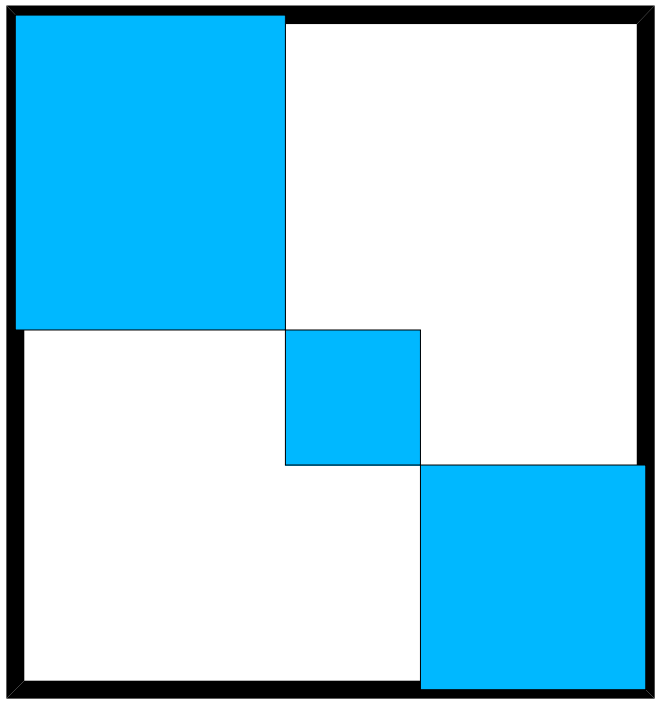
$$per(A)$$

Van der Waerden's Inequality

Two sided intervals $\in$ One sided intervals

Minc's Inequality

$$per(A) \leq \prod_{i=1}^{n} (r_i!)^{1/r_i} \approx \frac{r_i + 1/2 \log(2\pi r_i)}{e}$$



What would be nice...

$$M(\vec{r}) \leq M(\vec{r}^{1}) + \ldots + M(\vec{r}^{n})$$

				Row Sums
1	1	0	1	2
1	1	0	0	2
1	0	1	1	3
0	1	1	1	3

The Algorithm

For each column

Roll $U \in \{1,\ldots,M(\vec{r})\}$

Choose column i accordingly

Set $M(\vec{r}) \leftarrow M(\vec{r}^{\,i})$

Weak Minc's Inequality

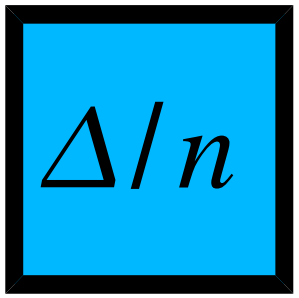
$$M'(\vec{r}) = \prod_{i=1}^{n} \frac{r_i + \frac{3+4\epsilon}{2\epsilon} + (1/2+\epsilon)\log\left(r_i + \frac{3+4\epsilon}{2}\epsilon\right)}{e}$$

$$\frac{M'(\vec{r}^{\,1}) + \ldots + M'(\vec{r}^{\,n})}{M'(\vec{r})} \leq 1$$

Find maximizing values of r_i
First show that r_i are equal
Then use Taylor series expansions

Lower Bound

Van der Waerden Inequality



Row sums, column sums Δ

Consequence:

$$per(A) \geq n! \frac{\Delta^n}{n^n} \approx \sqrt{2\pi n}\left(\frac{\Delta}{e}\right)^n$$

Results

$$\left(\frac{\Delta+\frac{3+4\epsilon}{2\epsilon}+(1/2+\epsilon)\log\left(\Delta+\frac{3+4\epsilon}{2}\epsilon\right)}{e}\right)^n \geq per(A) \geq \sqrt{n}\left(\frac{\Delta}{e}\right)^n$$

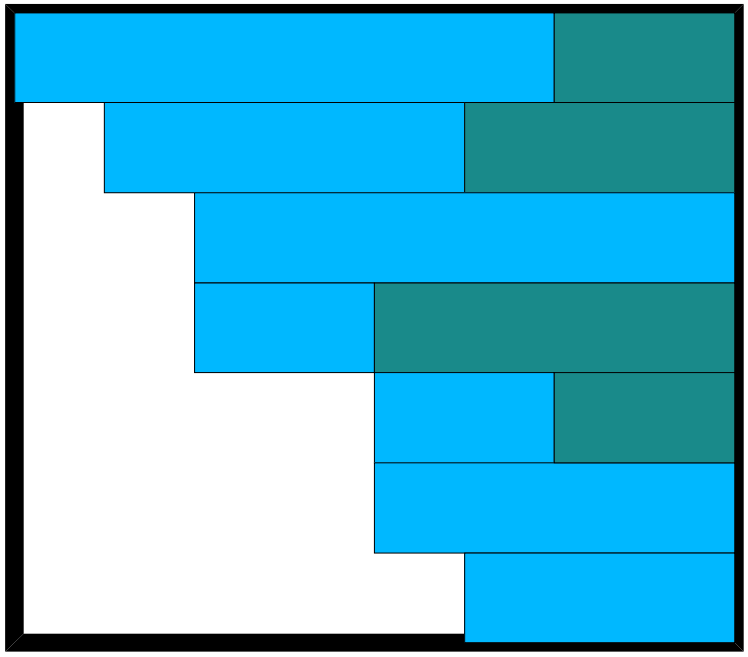
Running Time:

$$\frac{1}{\sqrt{n}}\exp\left[\left(\frac{3+4\epsilon}{2\epsilon}+(1/2+\epsilon)\log\left(\Delta+\frac{3+4\epsilon}{2}\epsilon\right)\right)\frac{n}{\Delta}\right]$$

$$\Delta = n\gamma \qquad O\left(n^{3/2+1/(2\gamma)+2/(\gamma\sqrt{\log n})}\right)$$

Second Approach

Two sided intervals $\in$ One sided intervals



Results

Conclusions

- Discrete acceptance/rejection for permanent
  - Upper/lower bounds
  - Need constructive proof
- Gives exact samples
  - No need to know mixing time
- Generation and Counting Easy