# Using Correlation Length to Compare MCMC Methods Graphically

Madeleine Thompson, advised by Radford Neal

April 29, 2010

# Motivation

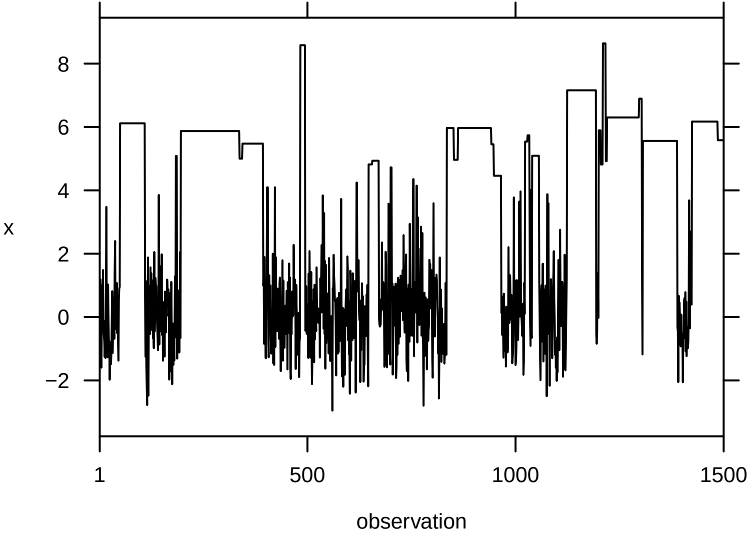
Context:

- MCMC methods construct a Markov chain of dependent samples from a target distribution
- Different methods work better on different distributions
- Extensive tuning may be required, limiting usefulness
- Comparisons between methods in existing research are often confusing

Goals:

- Present comparisons between MCMC methods clearly
- Minimize end-user tuning

# MCMC users spend a lot of time looking at trace plots



# Figures of merit

Two (usually equivalent) ways of describing how well an MCMC method performs on a specific distribution:

- Processor-seconds per independent observation
  - invariant to chain length
  - direct connection to user needs
  - but, depends on test hardware and system load
- Density function evaluations per independent observation
  - also invariant to chain length
  - does not depend on test hardware or system load
  - but, does not account for processor use by the sampler itself

But, what does "per independent observation" mean?

# Correlation length/autocorrelation time, $\tau$

- Define $\tau$ by:

$$\tau = 1 + 2\sum_{k=1}^{\infty} \rho_k \tag{1}$$

  where $\rho_k$ is the ACF of $\{X_j\}$ at lag $k$:

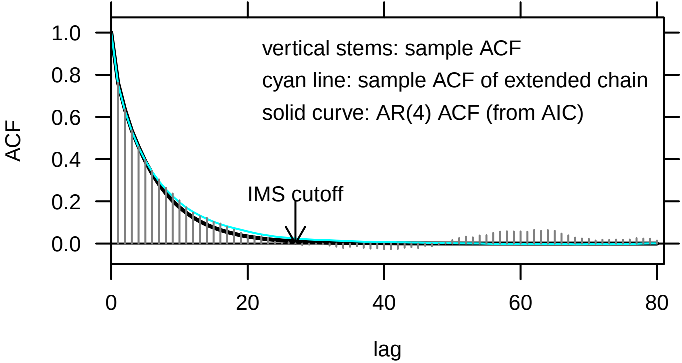
$$\rho_k = \frac{\text{cov}(X_j, X_{j-k})}{\text{var}(X_j)} \quad \text{(for all } j\text{)}$$

- If the sum in equation 1 converges, we have the CLT:

$$\sqrt{n/\tau}(\bar{X}_n - E(X_j)) \Rightarrow N\big(0, \text{var}(X_j)\big)$$

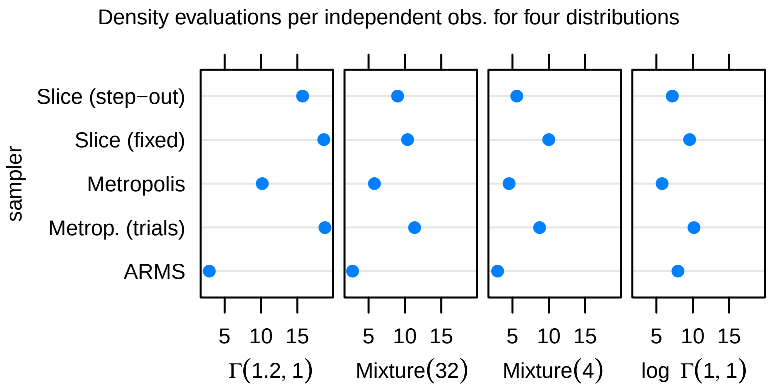
- The sample ACF is inaccurate for large lags, so we cannot use equation 1 to estimate $\tau$ directly.

# Two ways of modeling the ACF (of a difficult example)



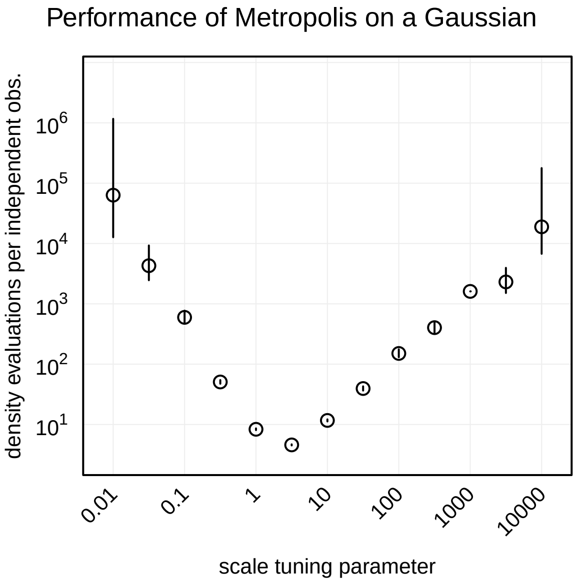

- Intial monotone sequence (IMS): sum sample ACF until a heuristic cutoff
- AR(AIC): autoregressive model with order chosen by AIC

# Comparing optimally-tuned samplers

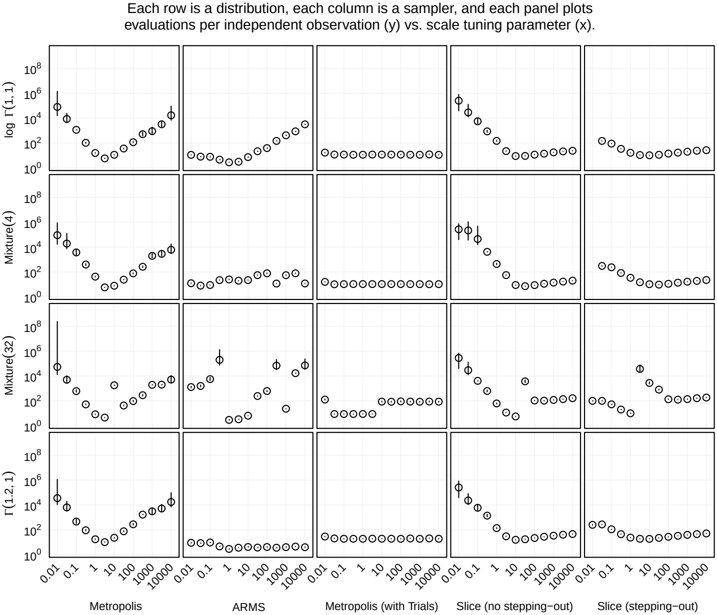



- Easy to read: ARMS appears to do well
- But, samplers are often not optimally tuned
- Narrow range of density evaluations per independent obs.

# A tuning parameter plot



# Comparing several samplers on several distributions



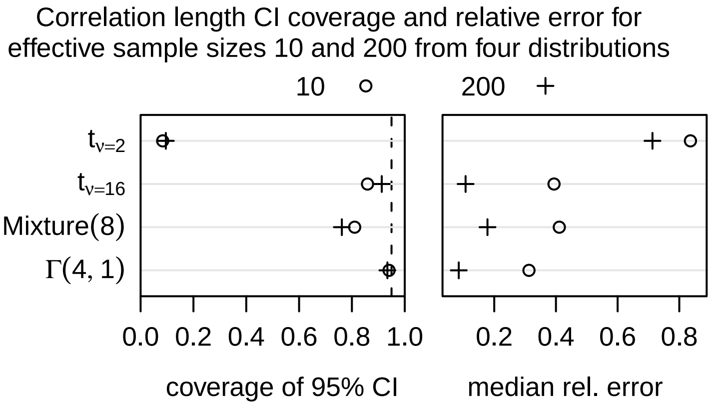
# Summary

- We can create grids of samplers and distributions plotting log-density evaluations per independent observation against a tuning parameter.
- Grids allow researchers to compare MCMC methods on a wide variety of distributions and tuning parameters.
- Comparison clarifies which methods are suitable for end-users with minimal knowledge of MCMC.

# References


- C. J. Geyer, "Practical Markov Chain Monte Carlo," *Statistical Science* 7 no. 4 (1992): 473–511.
- M. Plummer, N. Best, K. Cowles, and K. Vines, "CODA: Convergence Diagnosis and Output Analysis for MCMC," *R News* 6 no. 1 (Mar. 2006): 7–11.

# Performance of AR modeling of correlation length




- When target variance is defined (i.e. excepting $t_{\nu=2}$), nominal 95% CI for $\tau$ moderately underestimates true uncertainty
- Moderate relative errors allow broad comparisons with small effective sample sizes