An Estimate for the Variance of MCMC Estimates ($\hat{\mu}_n$).

Let $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_n, \ldots$ be a Markov chain with invariant distribution π.

Suppose we wish to estimate

$$\mu = \int_S g(\underline{x}) \, d\pi(\underline{x})$$

Under regularity assumptions [e.g. the M chain is irreducible and $E_\pi g^2(\underline{X}) < \infty$]

we have that

$$\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^{n} g(\underline{X}_i) \longrightarrow \mu \quad a.s. [\pi]$$

and

$$\sqrt{n}\,(\hat{\mu}_n - \mu) \xrightarrow{d} N(0, \sigma^2) \quad \text{where}$$

$$\sigma^2 = \mathrm{Var}_\pi\left(g(\underline{X}_1)\right) + 2\sum_{k=1}^{\infty} \mathrm{Cov}_\pi\left(g(\underline{X}_1), g(\underline{X}_{1+k})\right)$$

So,

$$\hat{\mu}_n \underset{\text{appr.}}{\sim} N\left(\mu, \frac{\sigma^2}{n}\right)$$

But the naive estimator:

$$\sqrt{\frac{1}{n(n-1)} \sum^{n} \left[g(\underline{X}_i) - \hat{\mu}_n\right]^2}$$

will not be a good estimate for the standard error of $\hat{\mu}_n$. Usually the variables $g(\underline{X}_1)$ and $g(\underline{X}_{1+k})$ are positively correlated.

The Batch Estimator for the SE($\hat{\mu}_n$)

Divide the chain in batches

$$\underline{X}_1, \dots, \underline{X}_b \longrightarrow \tilde{\mu}_1(b)$$

$$\underline{X}_{b+1}, \dots, \underline{X}_{2b} \longrightarrow \tilde{\mu}_2(b)$$

$$\underline{X}_{(k-1)b+1}, \dots, X_{kb} \longrightarrow \tilde{\mu}_k(b)$$

Clearly,

$$\boxed{\hat{\mu}_n = \frac{1}{k} \sum \tilde{\mu}_i(b)}$$

Moreover, if b is large enough

$$\sqrt{b}\,(\tilde{\mu}_i(b) - \mu) \sim N(0, \sigma^2)$$

$$\Rightarrow \quad \tilde{\mu}_i(b) \underset{\text{approx}}{\sim} N\left(\mu, \frac{\sigma^2}{b}\right)$$

Therefore

$$\frac{1}{(k-1)} \sum_{i=1}^{k} \left(\tilde{\mu}_i(b) - \hat{\mu}_n \right)^2 \quad \text{estimates} \quad \frac{\sigma^2}{b}$$

and

$$\frac{1}{(k-1)\,k} \sum_{i=1}^{k} \left(\tilde{\mu}_i(b) - \hat{\mu}_n \right)^2 \quad \text{estimates} \quad \frac{\sigma^2}{bk} = \boxed{\frac{\sigma^2}{n}}$$

<u>Note</u>: we are assuming that the batch means, $\tilde{\mu}_i(b)$ are approximately uncorrelated. This will be the case for "b" large enough.

Hence, the Standard Error estimate for $\hat{\mu}_n$ is

$$\boxed{SD(\hat{\mu}_n) = \sqrt{\frac{1}{k(k-1)} \sum_{i=1}^{k} \left[\tilde{\mu}_i(b) - \hat{\mu}_n \right]^2}}$$

<u>Note</u>: the autocorrelation of the $\tilde{\mu}_i(b)$ are computed for increasing values of b until it drops below some given $\epsilon \simeq 0.05$ or 0.01.