Model comparison using simulated posterior draws

Christopher A. Sims
Princeton University
sims@princeton.edu

October 20, 2013

The problem

- We have two or more models, indexed by i, each of which, with its prior, defines a joint pdf $p_i(y, \theta_i)$ for the data and the parameters.
- The posterior probabilities on the models are proportional to $s_i = \int p_i(y, \theta_i) d\theta_i$.
- We have no analytic formula for the integrals.

An identity that provides methods

If $q_1(\theta)$ and $q_2(\theta)$ are positive, integrable functions on the same domain (i.e. can be thought of as unnormalized probability densities), with $z_i = \int q_i(\theta) d\theta$, and if $\alpha(\theta)$ is any function such that $0 < \int \alpha(\theta) q_i(\theta) < \infty$, $i = 1, 2$, then

$$\frac{\int \frac{q_1(\theta) q_2(\theta) \alpha(\theta)}{z_2} \, d\theta}{\int \frac{q_2(\theta) q_1(\theta) \alpha(\theta)}{z_1} \, d\theta} = \frac{E_2[q_1 \alpha]}{E_1[q_2 \alpha]} = \frac{z_1}{z_2} \, .$$

Specific methods

importance sampling $\int q_2 = z_2 = 1$, $\alpha = 1/q_2$, $z_1 = E_2[q_1/q_2]$. Does not use MCMC draws. Blows up if q_1/q_2 is huge for some θ's.

modified harmonic mean $z_2 = 1$, $\alpha = 1/q_1$, $z_1 = 1/E_1[q2/q_1]$. Uses only MCMC draws. Blows up if q_1/q_2 is huge for some θ's.

bridge sampling Pick α so both $q_1\alpha$ and $q_2\alpha$ are bounded, e.g. $\alpha = 1/(q_1 + q_2)$. Uses draws from both q_1 and q_2.

Optimal α

- With same number of draws from q_1 and q_2, it's

$$\alpha = \frac{1}{z_1 q_2 + z_2 q_1}\,.$$

- Since we don't know z_1/z_2, this is not directly a help. But if our initial guess is off, we can update it and repeat — with new z_1/z_2, but re-using the old draws of q_2 and q_1.

Application of bridge sampling to finding a normalizing constant

i. Generate posterior draws $\{\theta_j\}$ using the posterior kernel $k(\theta)$ and MCMC.

ii. Pick a pdf f (so $\int f(\theta)\,d\theta = 1$) that is easy to draw from and generate a sample $\left\{\theta_j^*\right\}$ from it, probably by direct sampling. Ideally, $f()$ should be a good approximation to the normalized $k()$.

iii. Form

$$\frac{\sum_{j=1}^{N} \frac{k(\theta_j^*)}{f(\theta_j^*)+k(\theta_j^*)}}{\sum_{j=1}^{N} \frac{f(\theta_j)}{f(\theta_j)+k(\theta_j)}}.$$

If N is big enough, this will be a good estimator of $\int k(\theta)\,d\theta$.