Stat 543 2-21-05

Recall: Simulation-based approaches to Bayes inference (understanding of $g(\theta|x)$)

a) iid simulation from $g(\theta|x)$ (w/o knowing $\int L(\theta)g(\theta)\,d\theta$) rejection algorithm

b) MCMC simulation

- Gibbs Sampling
- Metropolis-Hastings

The goal is to produce simulated values

$\theta_1^*, \theta_2^*, \ldots$ whose empirical dsn

is guaranteed to approximate $g(\theta|x)$ so that

I can approximate

$$Q = \int q(\theta)\, g(\theta|x)\, d\theta$$

with $\frac{1}{n}\sum_{i=1}^{n} q(\theta_i^*)$

back to Metropolis-Hastings

- initial θ_0^*
- at step i specify

$$J_i(\theta'|\theta)$$

select a candidate

$$\theta_i^{**} \sim J_i(\cdot \mid \theta_{i-1}^*)$$

and calculate

$$r_i = \frac{L(\theta_i^{**})\, g(\theta_i^{**}) / J_i(\theta_i^{**} \mid \theta_{i-1}^*)}{L(\theta_{i-1}^*)\, g(\theta_{i-1}^*) / J_i(\theta_{i-1}^* \mid \theta_i^{**})}$$

Let $Y_i \sim$ Bernoulli $(\min(r_i, 1))$ and take

$$\theta_i^* = Y_i \theta_i^{**} + (1 - Y_i)\theta_{i-1}^*$$

Often, with wisely chosen J's this produces a sequence $\{\theta_i^*\}$ whose empirical dsns look like

$g(\theta|x)$ . . . ☺ all of this w/o knowing

$\int L(\theta) g(\theta) d\theta$

Skip 3.3 of B+D Minimax Thy

Some non-Bayesian Optimality Thy for SELE of $\gamma(\theta)$ – consider an estimator $\delta(X)$

$L(\theta, a) = (\theta - a)^2$

$$
\begin{aligned}
R(\theta, \delta) &= E_\theta \left(\delta(X) - \gamma(\theta) \right)^2 \\
&= MSE_\theta \left(\delta(X) \right) \\
&= Var_\theta\, \delta(X) + \underbrace{\left(E_\theta \delta(X) - \gamma(\theta) \right)^2}_{\text{bias}_\theta^2}
\end{aligned}
$$

Rao-Blackwell is about sufficiency + SELE — idea is that if $T(X)$ is sufficient I can always find something (an estimator) at least as good as $\delta(X)$ ... if $\delta(X)$ is silly the thing I find will be better

Thm (Rao-Blackwell) Suppose $\delta(X)$ is an estimator of $\gamma(\theta) \in \mathbb{R}^1$ with $E_\theta |\delta(X)| < \infty$ and that $T(X)$ is sufficient for θ. Let

$$\delta^*(t) = E\left[\delta(X) \mid T(X) = t\right]$$

then $\delta^*(T(X))$ is an estimator of $\gamma(\theta)$

with

$$MSE_\theta(\delta^*(T(X)) \le MSE_\theta(\delta(X))$$

$$\left\{\begin{array}{l} \text{and} \quad MSE_\theta(\delta^*(T(X)) < MSE_\theta(\delta(X)) \ \text{ if} \\ \theta \text{ is such that } \quad E_\theta(\delta(X))^2 < \infty \quad \text{and} \\ \qquad P_\theta[\delta(X) \neq \delta^*(T(X))] > 0 \end{array}\right\}$$

<u>Example</u> $X_1, X_2, \ldots, X_n$ iid Bernoulli p

$\gamma(p) = p \qquad T(X) = \sum X_i$ is sufficient

$\delta(X) = X_1 \qquad MSE_p(\delta(X)) = p(1-p)$

("clearly" $\frac{T(X)}{n}$ is better ... $MSE_p\left(\frac{T(X)}{n}\right) = \frac{p(1-p)}{n}$)

How to use Rao – Blackwell to improve on $\delta(X) = X_1$? Given that $T(x) = t$ the conditional dsn of $X \mid T(x) = t$ is uniform on

$$\left\{ x \in \{0,1\}^n \mid T(x) = t \right\}$$

~~so that given $T(x) = t$~~ a fraction $\frac{t}{n}$ of the elements of ~~this~~ have $x_1 = 1$ so

$$\delta^*(t) = E\left[\underbrace{X_1}_{\delta(X)} \mid T(X) = t\right] = 1\left(\frac{t}{n}\right) + 0\left(1 - \frac{t}{n}\right)$$

$$= \frac{t}{n}$$

my competing estimator (gotten from Rao-Blackwell construction!) is

$$\delta^*(T(X)) = \frac{T(X)}{n}$$ ☺

note this is better than $\delta(X)$ — I should have expected this ... as long as $p \in (0,1)$ there is positive p probability that

$$\frac{\sum X_i}{n} \neq X_1$$

and { } of Rao-Blackwell says that $\delta^*(T(X))$ will be better than $\delta(X)$ —

Note that if I had used X itself I would have gotten no improvement ... $T(X)$ is minimal sufficient