Lecture 23

- Metropolis-Hasting (E. Teller-1953)

given $X_i \sim \text{Poisson}(\lambda = a + bt_i)$
$K(a|bX_i) =$ " "
$K(b|aX_i) =$ " "

we need a way to sample an "a" and a "b" from conditionals to approx posterior.

- M-H Rejection Sampling

Main steps: draw a from $K(a|bX_i)$; b from $K(b|aX_i)$
* steps are identical, can't sample directly so: Substeps

sub step 1 create : $q(a_{t-1}, \phi)$ a transitional distribution
↑ other hyperparameters

draw "a" from $q(a_{t-1}, \phi)$
* substitute distribution; diffrerent from $P(a|bX_i)$
* now we must check if "a" is a good choice

sub step 2 M-H ratio

- $$r := \frac{\dfrac{P(a=a_t, b=b_{t-1} | X)}{f_q(a_t | a_{t-1}, \phi)}}{\dfrac{P(a=a_{t-1}, b=b_{t-1} | X)}{f_q(a_{t-1} | a_t, \phi)}}$$

ratio of posterior to transition if a_t is used

chance to go old → new

chance to go new → old

ratio of posterior to transition if reversed

$$\Rightarrow \frac{P(a=a_t, b=b_{t-1} | X)}{P(a=a_{t-1}, b=b_{t-1} | X)}$$

posterior of new data

posterior of old data

"Metroplis" alg

substep 3 Accept/Reject proposed a_t

accept only if $r \geq 1$

* M-H has been proven to converge
* M-H is most useful when sampling from a conditional Kernel

we only → if q is symmetric

* Note: Gibbs Sampler is special case of M-H

draw from $q(\theta_j, \gamma) = P(\theta_j \mid \theta_{-j})$ • the conditional is the Known transitional distribution

not including j

$$r := \frac{\left.\dfrac{P(\theta_j = \theta_{j,t} \mid \theta_{-j})}{P(\theta_j = \theta_{j,t} \mid \theta_{-j})}\right\} = 1}{\left.\dfrac{P(\theta_j = \theta_{j,t-1} \mid \theta_{-j})}{P(\theta_j = \theta_{j,t-1} \mid \theta_{-j})}\right\} = 1} = 1$$

• for $q = N(\theta_{t-1}, 1^2)$: $f_q(a_t \mid a_{t-1}, \phi) = f_q(a_{t-1} \mid a_t, \phi)$

the metropolis model/algorithm

• * So we use M-H (or Metropolis) when given a conditional kernel

Bayesian protocol:

① Collect data

② Pick $\mathcal{F}$: likelihood model

③ Pick $P(\theta)$: prior

These are ASSUMED

what if they are wrong?

④ Get posterior : use for inference

- compute
- Grid sample
- Gibbs sample
- M-H algorithm

* We model check $P(X \mid \theta)$ & $P(\theta)$ to ensure X makes sense

- Check #1 : passing Prior Predictive distribution check ?

$$P(X) = \int_{\Theta} P(X|\theta)\,P(\theta)\,d\theta$$

- does X make sense based on distribution $P(X)$
- X must fall in $P(X)$

- Check #2 : passing posterior replicate distribution check ?

$$P(X^*|X) = \int_{\Theta} P(X^*|\theta)\,P(\theta|X)\,d\theta$$

- can X come from $X^*|X$?
- does our model generate new X^* that resembles our actual data.