## Hawaiian Shirt: No!

**Told Byron:** Maybe

**Date:** Friday, February 8, 2019

Apology
post-exam
donut Friday!

## Putting the pieces together

**Data:** $y_1, \ldots, y_n \rightarrow$ model w/ $f(y|\theta)$

**Prior:** On $\theta \rightarrow \pi(\theta)$

Incorporating prior beliefs

**Posterior:** $$\boxed{\mathbb{P}(\theta|y) = \frac{\pi(\theta)\, f(y|\theta)}{\int \pi(\theta) f(y|\theta) d\theta}}$$

Now that we have chosen our models, we have to <u>fit</u> our posterior. (using Bayes rule).

But, this isn't always easy! Math is hard!

MCMC is used to sample from $\mathbb{P}(\theta|y)$ or to calculate expectations for functions of $\theta$ (e.g. profit from Lyft example).

"Accept with probability $\alpha$"

1) • Simulate $u \sim \text{Unif}(0,1)$

• Accept the sample if $u \le \alpha$

↳ why? b/c $\mathbb{P}(U \le \alpha) = \alpha$

2) • Simulate $B \sim \text{Bern}(\alpha)$

• Accept if $B == 1$

Example: (Rejection sampling)

Sample $\theta^{(1)} \sim g(\theta)$

$g$ = density of a $N(0,1)$

$g(\theta^{(1)}) \rightarrow$ density evaluated

$$= \frac{1}{\sqrt{2\pi}} e^{-(\theta^{(1)})^2/2}$$

M is chosen by you.

$\mathbb{P}(\theta^{(1)} | y)$ is also known because we know the functional form of $\mathbb{P}(\theta | y)$.

## Importance Sampling

* New situation: we do not know $\mathbb{P}(\theta|y)$ but would like to calculate expectations of $h(\theta)$ given $y$.

Recall $\mathbb{P}(\theta|y) = \dfrac{\pi(\theta) f(y|\theta)}{\int \pi(\theta) f(y|\theta) d\theta}$

↳ We don't know $\mathbb{P}(\theta|y)$ but we do know $q(\theta|y) = \overbrace{\pi(\theta) f(y|\theta)}^{\text{unnormalized density}}$

$h(\theta)$: function of $\theta$ that we'd like to know the expectation of.

Ex: $\text{var}(\theta|y) = \mathbb{E}[\underbrace{\theta^2}_{h_1(\theta)=\theta^2}|y] - \mathbb{E}[\underbrace{\theta}_{h_2(\theta)=\theta}|y]^2$

$g(\theta)$: Our "nice" proposal density function that we know how to simulate from.

Ex: $g(\theta) = 1$ $[U(0,1)]$

$q(\theta|y)$: Unnormalized density $= \pi(\theta) f(y|\theta)$

$\pi(\theta)$: Prior density for $\theta$

$f(y|\theta)$: Density of the data

$$\mathbb{P}(\theta|y) = \frac{q(\theta|y)}{\int q(\theta|y)\,d\theta}$$

$$\mathbb{E}[h(\theta)|y] = \int h(\theta)\,\mathbb{P}(\theta|y)\,d\theta$$

$$= \int h(\theta)\left(\frac{q(\theta|y)}{\int q(\theta|y)\,d\theta}\right)d\theta$$

since it integrates over all values of $\theta$ it can go out.

$$= \frac{\int h(\theta)\,q(\theta|y)\,d\theta}{\int q(\theta|y)\,d\theta}$$

Expectation from density we know $g(\theta)$ (Look @ side note)

$$= \frac{\int \left[\frac{h(\theta)\,q(\theta|y)}{g(\theta)}\right] g(\theta)\,d\theta}{\int \left[\frac{q(\theta|y)}{g(\theta)}\right] g(\theta)\,d\theta}$$

(multiply both numerator and dominator by $\frac{g(\theta)}{g(\theta)} = 1$)

Use Monte Carlo on this where $\theta^{(1)}, \dots, \theta^{(S)}$ will be drawn from $g(\theta)$.

= [Slide 19/23]

Side note:

1) $\theta \sim g(\theta)$

$$\mathbb{E}[\theta] = \int \theta \, g(\theta) \, d\theta$$

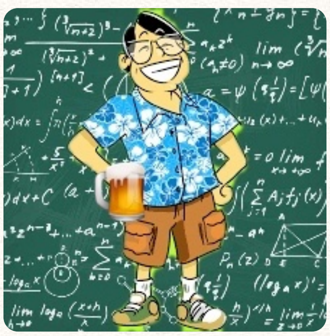
$$\mathbb{E}[\theta^2] = \int \theta^2 \, g(\theta) \, d\theta$$


2) $\theta \sim \mathbb{P}(\theta | y)$

$$\mathbb{E}[\theta] = \int \theta \, \mathbb{P}(\theta | y) \, d\theta$$

$$\mathbb{E}[\theta^2] = \int \theta^2 \, \mathbb{P}(\theta | y) \, d\theta$$

[R Script]



# Markov-Chain Simulation

* Big aim of MCMC → Sample from posteriors that we don't know.

1) The first MC (Markov Chain) gives us a way to sample iteratively where each sample depends on the last sample.

2) The sampling algorithm forms a Markov Chain, whose stationary distribution is the posterior.


~~Andrei Markov~~

Andrey Markov