# Monte Carlo Approximation

Jingchen (Monika) Hu

Vassar College

MATH 347 Bayesian Statistics

# Outline


1. Introduction
2. Procedure
3. Functions of $\theta$

# Outline


1. Introduction
2. Procedure
3. Functions of $\theta$

# Monte Carlo Approximation

Suppose we want to summarize posterior distribution of function of $\theta$, say $\phi = g(\theta)$. For example, we might want to compute the expectation, $E(\phi \mid Y)$.

- We have
$$E(\phi \mid Y) = \int_{g(\Theta)} \phi p(\phi \mid Y) d\phi = \int_{\Theta} g(\theta) p(\theta \mid Y) d\theta$$
- What if we do not know how to compute the integral?
- Common problem as we move in to higher dimensional parameters $(\theta_1, \theta_2, \ldots, \theta_p)$

Appeal to simulation and the Law of Large Numbers.

# Outline


1. Introduction
2. Procedure
3. Functions of $\theta$

# Simulation as Approximation

Suppose we can sample $S$ values from the posterior distribution of $\theta$, so that

$$\theta^{(1)}, \ldots, \theta^{(S)} \overset{\text{iid}}{\sim} \pi(\theta \mid Y)$$

- Law of Large Numbers

$$\begin{aligned} \mathrm{E}[\theta \mid Y] &\approx \frac{1}{S}\sum \theta^{(s)} \\ \mathrm{E}[g(\theta) \mid Y] &\approx \frac{1}{S}\sum g(\theta^{(s)}) \end{aligned}$$

Sample means converge to their expectations for large $S$.

# Simulated Distributions

$$\theta^{(1)}, \ldots, \theta^{(S)} \overset{\text{iid}}{\sim} \pi(\theta \mid Y)$$

- Cumulative ordered values approximate $F(\theta \mid Y)$ (empirical cdf)

$$P(\theta < c \mid Y) \approx \frac{\#(\theta^{(s)} < c)}{S}$$

how many number of draws out of S are less than c

- Empirical distribution of the sample $\theta^{(1)}, \ldots, \theta^{(S)}$ approximates $\pi(\theta \mid Y)$. Visualize with histogram or density estimator.
- Sample moments/quantiles/functions approximate true moments/quantiles/functions.
- For example, proportion of samples where event $g(\theta^{(s)}) > c$ approximates $P(g(\theta) > c \mid Y)$

Extends to higher dimensional parameters

# Tokyo Express Dining Preference Example

Posterior with Beta prior: $p \mid Y \sim \textrm{Beta}(15.06, 10.56)$ - 95% middle credible interval

- Exact solution: use the `beta_interval()` function in the `ProbBayes` package and or the `qbeta()` R function

```
beta_interval(0.95, c(15.06, 10.56), Color = crcblue)
```

```
c(qbeta(0.025, 15.06, 10.56), qbeta(0.975, 15.06, 10.56))
```

```
## [1] 0.3960802 0.7665697
```

# Tokyo Express Dining Preference Example cont'd

- Approximation through Monte Carlo simulation: use the `rbeta()` R function and `quantile()` R function ($S$ determines the accuracy. Make it larger when practical; usually 1000 is enough.)

```
S <- 1000; BetaSamples <- rbeta(S, 15.06, 10.56)
quantile(BetaSamples, c(0.025, 0.975))
```

```
##      2.5%     97.5%
## 0.3808316 0.7655909
```

# Outline


1. Introduction
2. Procedure
3. Functions of $\theta$

# Functions of $\theta$

Simulate posterior distribution of odds of preferring Friday:

$$o = \frac{p}{1-p} \Longrightarrow p = \frac{o}{1+o}, \quad \frac{dp}{do} = \frac{1}{(1+o)^2}$$

# Functions of $\theta$

Simulate posterior distribution of odds of preferring Friday:

$$o = \frac{p}{1-p} \Longrightarrow p = \frac{o}{1+o}, \quad \frac{dp}{do} = \frac{1}{(1+o)^2}$$

- Exact solution: change of variable $p \mid Y \sim \text{Beta}(a, b)$:

$$p(o \mid Y) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{o^{a-1}}{(1+o)^{a+b}}$$

# Functions of $\theta$

odds: ratio of success prob. over failure prob.

Simulate posterior distribution of odds of preferring Friday:

$$o = \frac{p}{1-p} \Longrightarrow p = \frac{o}{1+o}, \quad \frac{dp}{do} = \frac{1}{(1+o)^2}$$

- Exact solution: change of variable $p \mid Y \sim \text{Beta}(a, b)$:

$$p(o \mid Y) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{o^{a-1}}{(1+o)^{a+b}}$$

- Monte Carlo approximation: draw independent samples from $p(o \mid Y)$:

```
S <- 1000; BetaSamples <- rbeta(S, 15.06, 10.56)
odds = BetaSamples / (1 - BetaSamples)
```

# Monte Carlo Approximation to odds $o$

Monte Carlo approximation: draw independent samples from $p(o \mid Y)$:

```
S <- 1000; BetaSamples <- rbeta(S, 15.06, 10.56)
odds = BetaSamples / (1 - BetaSamples)
```

```
mean(odds)
```

```
## [1] 1.585919
```

```
median(odds)
```

```
## [1] 1.436583
```

```
quantile(odds, c(0.025, 0.975))
```

```
##      2.5%     97.5%
## 0.6720057 3.3439027
```

# Comparing Distributions

- Data from VA Hospitals: for each year observe $n$ patients and $y$, the number of cases (real failures).
- Observed data $Y = \{y_1, n_1;\ y_2, n_2\}$ for hospital 21:
  - In 1992, $y_1 = 306, n_1 = 651$
  - In 1993, $y_2 = 300, n_2 = 705$

# Comparing Distributions

- Data from VA Hospitals: for each year observe $n$ patients and $y$, the number of cases (real failures).
- Observed data $Y = \{y_1, n_1;\ y_2, n_2\}$ for hospital 21:
  - In 1992, $y_1 = 306, n_1 = 651$
  - In 1993, $y_2 = 300, n_2 = 705$
- First Model: Independent binomial outcomes in each year with probabilities $p_1$ and $p_2$.
- Question of Interest: has the probability changed between 1992 and 1993?

# Comparing Distributions cont'd

- Independent continuous Uniform priors $\rightarrow$ independent posteriors:
  - $\text{Uniform}(0, 1) = \text{Beta}(1, 1)$
  - In 1992, $y_1 = 306, n_1 = 651$
  - In 1993, $y_2 = 300, n_2 = 705$

# Comparing Distributions cont'd

- Independent continuous Uniform priors $\rightarrow$ independent posteriors:
  - $\mathrm{Uniform}(0,1) = \mathrm{Beta}(1,1)$
  - In 1992, $y_1 = 306, n_1 = 651$
  - In 1993, $y_2 = 300, n_2 = 705$
- $p_1 \mid Y \sim \mathrm{Beta}(307, 346)$ and
- $p_2 \mid Y \sim \mathrm{Beta}(301, 406)$ (independent of $\theta_1$)
- $p_i$ independent and $y_i \mid p_i$ independent imply $p_i$ independent a posterior

# Difference

New parameter $\delta = p_2 - p_1$ measures difference.

- Immediately:
  $E(\delta \mid Y) = E(p_2 \mid Y) - E(p_1 \mid Y) = 0.426 - 0.470 = -0.044.$
- Is this significantly different from 0? Is it really negative? (improvement in care)
- Immediately: $V(\delta \mid Y) = V(p_2 \mid Y) + V(p_1 \mid Y) = 0.0275^2$, sd $= 0.0275$
- mean $\pm$ 2 sd $= (-.044 \pm 2 \times 0.0275)$ includes zero (rough)

Can compute $p(\delta \mid Y)$ by transformation – but messy.

Use Monte Carlo Simulation!

# Posterior Simulation

Simulate large sample of $S$ values for $\theta_1$, similar for $\theta_2$ and then compute $\delta$

```
y1 <- 306; y2 <- 300; n1 <- 651; n2 <- 705;
S <- 5000;
t1 <- rbeta(S, y1 + 1, n1 - y1 + 1);
t2 <- rbeta(S, y2 + 1, n2 - y2 + 1);
d <- t2 - t1
sum(d < 0) / S
```

```
## [1] 0.9492
```

About a 95% posterior probability that $\delta < 0$