Lab 7 (part 2): Gibbs sampling

Olivier Binette

Friday October 9, 2020

Agenda

- Quick review of Gibbs sampling
- Homework: censoring problem
- Office hours

Quick review of Gibbs sampling

Quick review of Gibbs sampling

Given a likelihood $p(X \mid \theta_1, \theta_2)$ and a prior $p(\theta_1, \theta_2)$, you want to sample from the posterior $p(\theta_1, \theta_2 \mid X)$.

- You've seen techniques to sample from univariate distributions
  - Inverse CDF technique
  - Rejection sampling
- Now how do you sample from complicated **multivariate** distributions?

Quick review of Gibbs sampling

Given a likelihood $p(X \mid \theta_1, \theta_2)$ and a prior $p(\theta_1, \theta_2)$, you want to sample from the posterior $p(\theta_1, \theta_2 \mid X)$.

- You've seen techniques to sample from univariate distributions
  - Inverse CDF technique
  - Rejection sampling
- Now how do you sample from complicated **multivariate** distributions?

Quick review of Gibbs sampling

Given a likelihood $p(X \mid \theta_1, \theta_2)$ and a prior $p(\theta_1, \theta_2)$, you want to sample from the posterior $p(\theta_1, \theta_2 \mid X)$.

- You've seen techniques to sample from univariate distributions
  - Inverse CDF technique
  - Rejection sampling
- Now how do you sample from complicated **multivariate** distributions?

Quick review of Gibbs sampling

Given a likelihood $p(X \mid \theta_1, \theta_2)$ and a prior $p(\theta_1, \theta_2)$, you want to sample from the posterior $p(\theta_1, \theta_2 \mid X)$.

- You've seen techniques to sample from univariate distributions
  - Inverse CDF technique
  - Rejection sampling
- Now how do you sample from complicated **multivariate** distributions?

Quick review of Gibbs sampling

Given a likelihood $p(X \mid \theta_1, \theta_2)$ and a prior $p(\theta_1, \theta_2)$, you want to sample from the posterior $p(\theta_1, \theta_2 \mid X)$.

- You've seen techniques to sample from univariate distributions
  - Inverse CDF technique
  - Rejection sampling
- Now how do you sample from complicated **multivariate** distributions?

Quick review of Gibbs sampling

Problem: We want k samples $(\theta_1^{(s)}, \theta_2^{(s)})$, $s = 1, 2, \ldots, k$, from the posterior distribution $p(\theta_1, \theta_2 \mid X)$.

- The samples $(\theta_1^{(s)}, \theta_2^{(s)})$ may not be independent, and that's ok.
- We just want to be able to use the samples $(\theta_1^{(s)}, \theta_2^{(s)})$ to be able to approximate the posterior distribution.

Quick review of Gibbs sampling

Problem: We want k samples $(\theta_1^{(s)}, \theta_2^{(s)})$, $s = 1, 2, \ldots, k$, from the posterior distribution $p(\theta_1, \theta_2 \mid X)$.

- The samples $(\theta_1^{(s)}, \theta_2^{(s)})$ may not be independent, and that's ok.
- We just want to be able to use the samples $(\theta_1^{(s)}, \theta_2^{(s)})$ to be able to approximate the posterior distribution.

Quick review of Gibbs sampling

Problem: We want k samples $(\theta_1^{(s)}, \theta_2^{(s)})$, $s = 1, 2, \ldots, k$, from the posterior distribution $p(\theta_1, \theta_2 \mid X)$.

- The samples $(\theta_1^{(s)}, \theta_2^{(s)})$ may not be independent, and that's ok.
- We just want to be able to use the samples $(\theta_1^{(s)}, \theta_2^{(s)})$ to be able to approximate the posterior distribution.

Quick review of Gibbs sampling

Problem: We want k samples $(\theta_1^{(s)}, \theta_2^{(s)})$, $s = 1, 2, \ldots, k$, from the posterior distribution $p(\theta_1, \theta_2 \mid X)$.

- The samples $(\theta_1^{(s)}, \theta_2^{(s)})$ may not be independent, and that's ok.
- We just want to be able to use the samples $(\theta_1^{(s)}, \theta_2^{(s)})$ to be able to approximate the posterior distribution.

Quick review of Gibbs sampling

Gibbs sampling algorithm:

- For $s = 1$, initialize $\theta_1^{(1)}$ and $\theta_2^{(1)}$ to reasonable values.
- For $s = 2, 3, \ldots, k$, do:
  - $\theta_1^{(s)} \sim p(\theta_1 \mid X, \theta_2^{(s-1)})$,
  - $\theta_2^{(s)} \sim p(\theta_2 \mid X, \theta_1^{(s)})$.

Exercise: Compare to Lab 7 from last week.

Quick review of Gibbs sampling

Gibbs sampling algorithm:

- For $s = 1$, initialize $\theta_1^{(1)}$ and $\theta_2^{(1)}$ to reasonable values.
- For $s = 2, 3, \ldots, k$, do:
  - $\theta_1^{(s)} \sim p(\theta_1 \mid X, \theta_2^{(s-1)})$,
  - $\theta_2^{(s)} \sim p(\theta_2 \mid X, \theta_1^{(s)})$.

Exercise: Compare to Lab 7 from last week.

Quick review of Gibbs sampling

Gibbs sampling algorithm:

- For $s = 1$, initialize $\theta_1^{(1)}$ and $\theta_2^{(1)}$ to reasonable values.
- For $s = 2, 3, \ldots, k$, do:
  - $\theta_1^{(s)} \sim p(\theta_1 \mid X, \theta_2^{(s-1)})$,
  - $\theta_2^{(s)} \sim p(\theta_2 \mid X, \theta_1^{(s)})$.

Exercise: Compare to Lab 7 from last week.

Quick review of Gibbs sampling

Gibbs sampling algorithm:

- For $s = 1$, initialize $\theta_1^{(1)}$ and $\theta_2^{(1)}$ to reasonable values.
- For $s = 2, 3, \ldots, k$, do:
  - $\theta_1^{(s)} \sim p(\theta_1 \mid X, \theta_2^{(s-1)})$,
  - $\theta_2^{(s)} \sim p(\theta_2 \mid X, \theta_1^{(s)})$.

Exercise: Compare to Lab 7 from last week.

Quick review of Gibbs sampling

Gibbs sampling algorithm:

- For $s = 1$, initialize $\theta_1^{(1)}$ and $\theta_2^{(1)}$ to reasonable values.
- For $s = 2, 3, \ldots, k$, do:
  - $\theta_1^{(s)} \sim p(\theta_1 \mid X, \theta_2^{(s-1)})$,
  - $\theta_2^{(s)} \sim p(\theta_2 \mid X, \theta_1^{(s)})$.

Exercise: Compare to Lab 7 from last week.

Quick review of Gibbs sampling

Gibbs sampling algorithm:

- For $s = 1$, initialize $\theta_1^{(1)}$ and $\theta_2^{(1)}$ to reasonable values.
- For $s = 2, 3, \ldots, k$, do:
  - $\theta_1^{(s)} \sim p(\theta_1 \mid X, \theta_2^{(s-1)})$,
  - $\theta_2^{(s)} \sim p(\theta_2 \mid X, \theta_1^{(s)})$.

Exercise: Compare to Lab 7 from last week.

Quick review of Gibbs sampling

What's great about Gibbs sampling?

- At every step, you're only sampling from a univariate distribution.
- Sampling from a univariate distribution (the full conditionals) is (almost always) doable.
  - There are even "black box" algorithms for that.
  - See Luc Devroye's "Non-Uniform Random Variate Generation" (1986) if interested.

Quick review of Gibbs sampling

What's great about Gibbs sampling?

- At every step, you're only sampling from a univariate distribution.
- Sampling from a univariate distribution (the full conditionals) is (almost always) doable.
  - There are even "black box" algorithms for that.
  - See Luc Devroye's "Non-Uniform Random Variate Generation" (1986) if interested.

Quick review of Gibbs sampling

What's great about Gibbs sampling?

- At every step, you're only sampling from a univariate distribution.
- Sampling from a univariate distribution (the full conditionals) is (almost always) doable.
  - There are even "black box" algorithms for that.
  - See Luc Devroye's "Non-Uniform Random Variate Generation" (1986) if interested.

Quick review of Gibbs sampling

What's great about Gibbs sampling?

- At every step, you're only sampling from a univariate distribution.
- Sampling from a univariate distribution (the full conditionals) is (almost always) doable.
  - There are even "black box" algorithms for that.
  - See Luc Devroye's "Non-Uniform Random Variate Generation" (1986) if interested.

Quick review of Gibbs sampling

What's great about Gibbs sampling?

- At every step, you're only sampling from a univariate distribution.
- Sampling from a univariate distribution (the full conditionals) is (almost always) doable.
  - There are even "black box" algorithms for that.
  - See Luc Devroye's "Non-Uniform Random Variate Generation" (1986) if interested.

Quick review of Gibbs sampling

What's great about Gibbs sampling?

- At every step, you're only sampling from a univariate distribution.
- Sampling from a univariate distribution (the full conditionals) is (almost always) doable.
  - There are even "black box" algorithms for that.
  - See Luc Devroye's "Non-Uniform Random Variate Generation" (1986) if interested.

Homework: censoring problem

Homework: censoring problem

- You're interested in the survival times $Z_1, \ldots, Z_n$ of individuals following a medical procedure.
- You model the Z_i as being i.i.d. Gamma(r, θ)
- r is known and you have a prior $\theta \sim$ Gamma(a, b).

Unfortunately, you only observe the variables X_i defined as:

- $X_i = Z_i$ if $Z_i < c_i$;
- $X_i = c_i$ if $Z_i > c_i$.

The censoring times c_i are given to you when there is censoring (they don't matter when there's no censoring).

Homework: censoring problem

- You're interested in the survival times $Z_1, \ldots, Z_n$ of individuals following a medical procedure.
- You model the Z_i as being i.i.d. Gamma(r, θ)
- r is known and you have a prior $\theta \sim$ Gamma(a, b).

Unfortunately, you only observe the variables X_i defined as:

- $X_i = Z_i$ if $Z_i < c_i$;
- $X_i = c_i$ if $Z_i > c_i$.

The censoring times c_i are given to you when there is censoring (they don't matter when there's no censoring).

Homework: censoring problem

- You're interested in the survival times $Z_1, \ldots, Z_n$ of individuals following a medical procedure.
- You model the Z_i as being i.i.d. Gamma(r, θ)
- r is known and you have a prior $\theta \sim$ Gamma(a, b).

Unfortunately, you only observe the variables X_i defined as:

- $X_i = Z_i$ if $Z_i < c_i$;
- $X_i = c_i$ if $Z_i > c_i$.

The censoring times c_i are given to you when there is censoring (they don't matter when there's no censoring).

Homework: censoring problem

- You're interested in the survival times $Z_1, \ldots, Z_n$ of individuals following a medical procedure.
- You model the Z_i as being i.i.d. Gamma(r, θ)
- r is known and you have a prior $\theta \sim$ Gamma(a, b).

Unfortunately, you only observe the variables X_i defined as:

- $X_i = Z_i$ if $Z_i < c_i$;
- $X_i = c_i$ if $Z_i > c_i$.

The censoring times c_i are given to you when there is censoring (they don't matter when there's no censoring).

Homework: censoring problem

The **goal** is to do inference on θ.

$$
\begin{aligned}
p(\theta \mid x_{1:n}) &\propto p(x_{1:n} \mid \theta)p(\theta) \\
&= \int p(x_{1:n}, z_{1:n} \mid \theta)\, dz_{1:n}\, p(\theta) \\
&= \int p(x_{1:n} \mid z_{1:n}, \theta)p(z_{1:n} \mid \theta)\, dz_{1:n}\, p(\theta)
\end{aligned}
$$

where

$$
p(x_{1:n} \mid z_{1:n}, \theta) = \prod_{i=1}^{n} \left(\mathbb{I}(x_i = z_i)\mathbb{I}(z_i < c_i) + \mathbb{I}(x_i = c_i)(z_i > c_i)\right)
$$

and

$$
p(z_{1:n} \mid \theta) = \prod_{i=1}^{n} \text{Gamma}(z_i; r, \theta).
$$

Homework: censoring problem

The **goal** is to do inference on θ.

$$
\begin{aligned}
p(\theta \mid x_{1:n}) &\propto p(x_{1:n} \mid \theta) p(\theta) \\
&= \int p(x_{1:n}, z_{1:n} \mid \theta) \, dz_{1:n} \, p(\theta) \\
&= \int p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) \, dz_{1:n} \, p(\theta)
\end{aligned}
$$

where

$$
p(x_{1:n} \mid z_{1:n}, \theta) = \prod_{i=1}^{n} \left(\mathbb{I}(x_i = z_i) \mathbb{I}(z_i < c_i) + \mathbb{I}(x_i = c_i)(z_i > c_i) \right)
$$

and

$$
p(z_{1:n} \mid \theta) = \prod_{i=1}^{n} \mathrm{Gamma}(z_i; r, \theta).
$$

Homework: censoring problem

The **goal** is to do inference on θ.

$$\begin{aligned}
p(\theta \mid x_{1:n}) &\propto p(x_{1:n} \mid \theta) p(\theta) \\
&= \int p(x_{1:n}, z_{1:n} \mid \theta)\, dz_{1:n}\, p(\theta) \\
&= \int p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta)\, dz_{1:n}\, p(\theta)
\end{aligned}$$

where

$$p(x_{1:n} \mid z_{1:n}, \theta) = \prod_{i=1}^{n} \left(\mathbb{I}(x_i = z_i)\mathbb{I}(z_i < c_i) + \mathbb{I}(x_i = c_i)(z_i > c_i) \right)$$

and

$$p(z_{1:n} \mid \theta) = \prod_{i=1}^{n} \text{Gamma}(z_i; r, \theta).$$

Homework: censoring problem

The **goal** is to do inference on θ.

$$\begin{aligned} p(\theta \mid x_{1:n}) &\propto p(x_{1:n} \mid \theta) p(\theta) \\ &= \int p(x_{1:n}, z_{1:n} \mid \theta)\, dz_{1:n}\, p(\theta) \\ &= \int p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta)\, dz_{1:n}\, p(\theta) \end{aligned}$$

where

$$p(x_{1:n} \mid z_{1:n}, \theta) = \prod_{i=1}^{n} \left(\mathbb{I}(x_i = z_i)\mathbb{I}(z_i < c_i) + \mathbb{I}(x_i = c_i)(z_i > c_i) \right)$$

and

$$p(z_{1:n} \mid \theta) = \prod_{i=1}^{n} \text{Gamma}(z_i; r, \theta).$$

Homework: censoring problem

The **goal** is to do inference on θ.

$$
\begin{aligned}
p(\theta \mid x_{1:n}) &\propto p(x_{1:n} \mid \theta) p(\theta) \\
&= \int p(x_{1:n}, z_{1:n} \mid \theta)\, dz_{1:n}\, p(\theta) \\
&= \int p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta)\, dz_{1:n}\, p(\theta)
\end{aligned}
$$

where

$$
p(x_{1:n} \mid z_{1:n}, \theta) = \prod_{i=1}^{n} \left(\mathbb{I}(x_i = z_i)\mathbb{I}(z_i < c_i) + \mathbb{I}(x_i = c_i)(z_i > c_i) \right)
$$

and

$$
p(z_{1:n} \mid \theta) = \prod_{i=1}^{n} \mathrm{Gamma}(z_i; r, \theta).
$$

Homework: censoring problem

The **goal** is to do inference on θ.

$$\begin{aligned} p(\theta \mid x_{1:n}) &\propto p(x_{1:n} \mid \theta) p(\theta) \\ &= \int p(x_{1:n}, z_{1:n} \mid \theta)\, dz_{1:n}\, p(\theta) \\ &= \int p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta)\, dz_{1:n}\, p(\theta) \end{aligned}$$

where

$$p(x_{1:n} \mid z_{1:n}, \theta) = \prod_{i=1}^{n} \left(\mathbb{I}(x_i = z_i) \mathbb{I}(z_i < c_i) + \mathbb{I}(x_i = c_i)(z_i > c_i) \right)$$

and

$$p(z_{1:n} \mid \theta) = \prod_{i=1}^{n} \text{Gamma}(z_i; r, \theta).$$

Homework: censoring problem

Well that's a bit hard to compute...

Alternative: Let's sample from the joint posterior $p(\theta, z_{1:n} \mid x_{1:n})$ and then we can forget about $z_{1:n}$.

How do we do that? Gibbs sampling!

Homework: censoring problem

Well that's a bit hard to compute...

Alternative: Let's sample from the joint posterior $p(\theta, z_{1:n} \mid x_{1:n})$ and then we can forget about $z_{1:n}$.

How do we do that? Gibbs sampling!

Homework: censoring problem

Well that's a bit hard to compute...

Alternative: Let's sample from the joint posterior $p(\theta, z_{1:n} \mid x_{1:n})$ and then we can forget about $z_{1:n}$.

How do we do that? Gibbs sampling!

Homework: censoring problem

Well that's a bit hard to compute. . .

Alternative: Let's sample from the joint posterior $p(\theta, z_{1:n} \mid x_{1:n})$ and then we can forget about $z_{1:n}$.

How do we do that? Gibbs sampling!

Homework: censoring problem

We want to sample from $p(\theta, z_{1:n} \mid x_{1:n})$ using Gibbs sampling.

Step 1: Write down the main chunks of the full joint distribution.

$$p(\theta, z_{1:n}, x_{1:n}) = p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta)$$

Step 2: Derive the full conditional distributions.

$$\begin{aligned}
p(\theta \mid z_{1:n}, x_{1:n}) &\propto p(\theta, z_{1:n}, x_{1:n}) \\
&= p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta) \\
&\propto p(z_{1:n} \mid \theta) p(\theta) \\
&\propto \left(\prod_{i=1}^{n} \theta^r e^{-\theta z_i} \right) \theta^{a-1} e^{-b\theta} \\
&= \theta^{nr+a-1} e^{-(b+\sum_i z_i)\theta} \\
&\Rightarrow \theta \mid z_{1:n}, x_{1:n} \sim \text{Gamma}(nr + a, b + \sum_i z_i).
\end{aligned}$$

Homework: censoring problem

We want to sample from $p(\theta, z_{1:n} \mid x_{1:n})$ using Gibbs sampling.

Step 1: Write down the main chunks of the full joint distribution.

$$p(\theta, z_{1:n}, x_{1:n}) = p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta)$$

Step 2: Derive the full conditional distributions.

$$\begin{aligned}
p(\theta \mid z_{1:n}, x_{1:n}) &\propto p(\theta, z_{1:n}, x_{1:n}) \\
&= p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta) \\
&\propto p(z_{1:n} \mid \theta) p(\theta) \\
&\propto \left(\prod_{i=1}^{n} \theta^r e^{-\theta z_i} \right) \theta^{a-1} e^{-b\theta} \\
&= \theta^{nr+a-1} e^{-(b + \sum_i Z_i)\theta} \\
&\Rightarrow \theta \mid z_{1:n}, x_{1:n} \sim \text{Gamma}(nr + a, b + \sum_i Z_i).
\end{aligned}$$

Homework: censoring problem

We want to sample from $p(\theta, z_{1:n} \mid x_{1:n})$ using Gibbs sampling.

Step 1: Write down the main chunks of the full joint distribution.

$$p(\theta, z_{1:n}, x_{1:n}) = p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta)$$

Step 2: Derive the full conditional distributions.

$$\begin{aligned}
p(\theta \mid z_{1:n}, x_{1:n}) &\propto p(\theta, z_{1:n}, x_{1:n}) \\
&= p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta) \\
&\propto p(z_{1:n} \mid \theta) p(\theta) \\
&\propto \left(\prod_{i=1}^{n} \theta^r e^{-\theta z_i} \right) \theta^{a-1} e^{-b\theta} \\
&= \theta^{nr+a-1} e^{-(b+\sum_i z_i)\theta} \\
&\Rightarrow \theta \mid z_{1:n}, x_{1:n} \sim \text{Gamma}(nr + a, b + \sum_i z_i).
\end{aligned}$$

Homework: censoring problem

We want to sample from $p(\theta, z_{1:n} \mid x_{1:n})$ using Gibbs sampling.

Step 1: Write down the main chunks of the full joint distribution.

$$p(\theta, z_{1:n}, x_{1:n}) = p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta)$$

Step 2: Derive the full conditional distributions.

$$\begin{aligned}
p(\theta \mid z_{1:n}, x_{1:n}) &\propto p(\theta, z_{1:n}, x_{1:n}) \\
&= p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta) \\
&\propto p(z_{1:n} \mid \theta) p(\theta) \\
&\propto \left(\prod_{i=1}^{n} \theta^r e^{-\theta z_i} \right) \theta^{a-1} e^{-b\theta} \\
&= \theta^{nr+a-1} e^{-(b+\sum_i z_i)\theta} \\
&\Rightarrow \theta \mid z_{1:n}, x_{1:n} \sim \text{Gamma}(nr + a, b + \sum_i z_i).
\end{aligned}$$

Homework: censoring problem

We want to sample from $p(\theta, z_{1:n} \mid x_{1:n})$ using Gibbs sampling.

Step 1: Write down the main chunks of the full joint distribution.

$$p(\theta, z_{1:n}, x_{1:n}) = p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta)$$

Step 2: Derive the full conditional distributions.

$$\begin{aligned}
p(\theta \mid z_{1:n}, x_{1:n}) &\propto p(\theta, z_{1:n}, x_{1:n}) \\
&= p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta) \\
&\propto p(z_{1:n} \mid \theta) p(\theta) \\
&\propto \left(\prod_{i=1}^{n} \theta^{r} e^{-\theta z_i} \right) \theta^{a-1} e^{-b\theta} \\
&= \theta^{nr+a-1} e^{-(b+\sum_i Z_i)\theta} \\
&\Rightarrow \theta \mid z_{1:n}, x_{1:n} \sim \mathrm{Gamma}(nr + a, b + \sum_i Z_i).
\end{aligned}$$

Homework: censoring problem

We want to sample from $p(\theta, z_{1:n} \mid x_{1:n})$ using Gibbs sampling.

Step 1: Write down the main chunks of the full joint distribution.

$$p(\theta, z_{1:n}, x_{1:n}) = p(x_{1:n} \mid z_{1:n}, \theta)p(z_{1:n} \mid \theta)p(\theta)$$

Step 2: Derive the full conditional distributions.

$$\begin{aligned}
p(\theta \mid z_{1:n}, x_{1:n}) &\propto p(\theta, z_{1:n}, x_{1:n}) \\
&= p(x_{1:n} \mid z_{1:n}, \theta)p(z_{1:n} \mid \theta)p(\theta) \\
&\propto p(z_{1:n} \mid \theta)p(\theta) \\
&\propto \left(\prod_{i=1}^{n} \theta^r e^{-\theta z_i}\right) \theta^{a-1} e^{-b\theta} \\
&= \theta^{nr+a-1} e^{-(b+\sum_i z_i)\theta} \\
&\Rightarrow \theta \mid z_{1:n}, x_{1:n} \sim \text{Gamma}(nr+a, b+\sum_i z_i).
\end{aligned}$$

Homework: censoring problem

We want to sample from $p(\theta, z_{1:n} \mid x_{1:n})$ using Gibbs sampling.

Step 1: Write down the main chunks of the full joint distribution.

$$p(\theta, z_{1:n}, x_{1:n}) = p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta)$$

Step 2: Derive the full conditional distributions.

$$\begin{aligned}
p(\theta \mid z_{1:n}, x_{1:n}) &\propto p(\theta, z_{1:n}, x_{1:n}) \\
&= p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta) \\
&\propto p(z_{1:n} \mid \theta) p(\theta) \\
&\propto \left(\prod_{i=1}^{n} \theta^r e^{-\theta z_i} \right) \theta^{a-1} e^{-b\theta} \\
&= \theta^{nr+a-1} e^{-(b+\sum_i Z_i)\theta} \\
&\Rightarrow \theta \mid z_{1:n}, x_{1:n} \sim \text{Gamma}(nr + a, b + \sum_i Z_i).
\end{aligned}$$

Homework: censoring problem

The next full conditionals:

$$
\begin{aligned}
p(z_i \mid z_{(-i)}, x_{1:n}, \theta) &\propto p(\theta, z_{1:n}, x_{1:n}) \\
&= p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta) \\
&\propto p(x_i \mid z_i, \theta) p(z_i \mid \theta) \\
&\propto \begin{cases} \mathbb{I}(z_i = x_i) & \text{if } x_i \neq c_i \\ \mathbb{I}(z_i >= c_i) \text{Gamma}(z_i; r, \theta) & \text{if } x_i = c_i \end{cases}
\end{aligned}
$$

Beka gave you a function to sample from the truncated Gamma.

Homework: censoring problem

The next full conditionals:

$$
\begin{aligned}
p(z_i \mid z_{(-i)}, x_{1:n}, \theta) &\propto p(\theta, z_{1:n}, x_{1:n}) \\
&= p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta) \\
&\propto p(x_i \mid z_i, \theta) p(z_i \mid \theta) \\
&\propto \begin{cases} \mathbb{I}(z_i = x_i) & \text{if } x_i \neq c_i \\ \mathbb{I}(z_i >= c_i)\mathrm{Gamma}(z_i; r, \theta) & \text{if } x_i = c_i \end{cases}
\end{aligned}
$$

Beka gave you a function to sample from the truncated Gamma.

Homework: censoring problem

The next full conditionals:

$$
\begin{aligned}
p(z_i \mid z_{(-i)}, x_{1:n}, \theta) &\propto p(\theta, z_{1:n}, x_{1:n}) \\
&= p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta) \\
&\propto p(x_i \mid z_i, \theta) p(z_i \mid \theta) \\
&\propto \begin{cases} \mathbb{I}(z_i = x_i) & \text{if } x_i \neq c_i \\ \mathbb{I}(z_i >= c_i)\text{Gamma}(z_i; r, \theta) & \text{if } x_i = c_i \end{cases}
\end{aligned}
$$

Beka gave you a function to sample from the truncated Gamma.

Homework: censoring problem

The next full conditionals:

$$
\begin{aligned}
p(z_i \mid z_{(-i)}, x_{1:n}, \theta) &\propto p(\theta, z_{1:n}, x_{1:n}) \\
&= p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta) \\
&\propto p(x_i \mid z_i, \theta) p(z_i \mid \theta) \\
&\propto \begin{cases} \mathbb{I}(z_i = x_i) & \text{if } x_i \neq c_i \\ \mathbb{I}(z_i >= c_i)\text{Gamma}(z_i; r, \theta) & \text{if } x_i = c_i \end{cases}
\end{aligned}
$$

Beka gave you a function to sample from the truncated Gamma.

Homework: censoring problem

The next full conditionals:

$$
\begin{aligned}
p(z_i \mid z_{(-i)}, x_{1:n}, \theta) &\propto p(\theta, z_{1:n}, x_{1:n}) \\
&= p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta) \\
&\propto p(x_i \mid z_i, \theta) p(z_i \mid \theta) \\
&\propto \begin{cases} \mathbb{I}(z_i = x_i) & \text{if } x_i \neq c_i \\ \mathbb{I}(z_i >= c_i) \text{Gamma}(z_i; r, \theta) & \text{if } x_i = c_i \end{cases}
\end{aligned}
$$

Beka gave you a function to sample from the truncated Gamma.

Homework: censoring problem

The next full conditionals:

$$
\begin{aligned}
p(z_i \mid z_{(-i)}, x_{1:n}, \theta) &\propto p(\theta, z_{1:n}, x_{1:n}) \\
&= p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta) \\
&\propto p(x_i \mid z_i, \theta) p(z_i \mid \theta) \\
&\propto \begin{cases} \mathbb{I}(z_i = x_i) & \text{if } x_i \neq c_i \\ \mathbb{I}(z_i >= c_i)\mathrm{Gamma}(z_i; r, \theta) & \text{if } x_i = c_i \end{cases}
\end{aligned}
$$

Beka gave you a function to sample from the truncated Gamma.

Homework: censoring problem

The next full conditionals:

$$
\begin{aligned}
p(z_i \mid z_{(-i)}, x_{1:n}, \theta) &\propto p(\theta, z_{1:n}, x_{1:n}) \\
&= p(x_{1:n} \mid z_{1:n}, \theta) p(z_{1:n} \mid \theta) p(\theta) \\
&\propto p(x_i \mid z_i, \theta) p(z_i \mid \theta) \\
&\propto \begin{cases} \mathbb{I}(z_i = x_i) & \text{if } x_i \neq c_i \\ \mathbb{I}(z_i >= c_i)\text{Gamma}(z_i; r, \theta) & \text{if } x_i = c_i \end{cases}
\end{aligned}
$$

Beka gave you a function to sample from the truncated Gamma.

Homework: censoring problem

Step 3: Iterate through sampling from the full conditionals.

- Beka gave you a function that does this.
- Make sure you understand how it works.
- Note that we only need to sample from the values Z_i which have been censored, and not from those that have been observed already.

Homework: censoring problem

Step 3: Iterate through sampling from the full conditionals.

- Beka gave you a function that does this.
- Make sure you understand how it works.
- Note that we only need to sample from the values Z_i which have been censored, and not from those that have been observed already.

Homework: censoring problem

Step 3: Iterate through sampling from the full conditionals.

- Beka gave you a function that does this.
- Make sure you understand how it works.
- Note that we only need to sample from the values Z_i which have been censored, and not from those that have been observed already.

Homework: censoring problem

Step 3: Iterate through sampling from the full conditionals.

- Beka gave you a function that does this.
- Make sure you understand how it works.
- Note that we only need to sample from the values Z_i which have been censored, and not from those that have been observed already.