

# INTRODUCTION TO BAYESIAN STATISTICS

CARSON JAMES

## CONTENTS

1. Introduction	1
2. Sampling	1
2.1. Inverse CDF Sampling	1
2.2. Conditional Chain Sampling	1
2.3. Importance Sampling	2
2.4. Rejection Sampling	2
3. Posterior Consistency	4
3.1. Introduction	4

## 1. INTRODUCTION

**Definition 1.0.1.** Let  $A \in \mathcal{B}(R^d)$  and  $\Theta \neq \emptyset$ . Suppose that  $m(A) > 0$ . We define

$$\mathcal{D}(A) = \{f \in L^1(A) : f \geq 0 \text{ and } \|f\|_1 = 1\}$$

and for  $\theta \in \Theta$ , we define

$$\mathcal{D}(A|\theta) = \{f : A \times \Theta \rightarrow \mathbb{R} : f(\cdot|\theta) \in \mathcal{D}(A)\}$$

## 2. SAMPLING

### 2.1. Inverse CDF Sampling.

### 2.2. Conditional Chain Sampling.

**Definition 2.2.1.** Let  $A \subset \mathbb{R}^d$  be open,  $a = (a_1, \dots, a_d) \in A$ . Define  $\tau_1 : \mathbb{R} \rightarrow \mathbb{R}^d$  and  $A_1 \subset \mathbb{R}$  by

- $\tau_1(x) = (x, a_2, \dots, a_d)$
- $A_1 = \tau_1^{-1}(A)$

Choose  $f_1 \in \mathcal{D}(A_1)$  and sample  $b_1 \sim f_1$ . For  $j \in \{2, \dots, d\}$ , define  $\tau_j : \mathbb{R} \rightarrow \mathbb{R}^n$ ,  $A_j$ , choose  $f_j$  and sample  $b_j$  inductively by

- setting  $\tau_j(x) = (b_1, \dots, b_{j-1}, x, a_{j+1}, \dots, a_d)$
- setting  $A_j = \tau_j^{-1}(A)$
- choosing  $f_j \in \mathcal{D}(A_j|b_1, \dots, b_{j-1})$  and setting  $b_j \sim f_j$

Note that  $\tau_j$  is continuous which implies that  $A_j = \tau_j^{-1}(A)$  is open.

**Exercise 2.2.2.** Let  $A \subset \mathbb{R}$  be open and  $a = (a_1, \dots, a_d) \in A$ . Define  $A_j$ ,  $f_j$  and  $b_j$  as above and define  $b \in A$  and  $f : A \rightarrow \mathbb{R}$  by

$$b = (b_1, \dots, b_d)$$

and

$$f(x_1, \dots, x_d) = \prod_{j=1}^n f_j(x_j)$$

Then

- (1)  $f \in \mathcal{D}(A)$
- (2)  $b \sim f$

*Proof.* (1) Fubini's theorem implies that

$$\begin{aligned} \int_A f dm^d &= \int f_1(x_1) \left[ \int f_2(x_2) \left[ \dots \left[ \int f_d(x_d) dm(x_d) \right] \dots \right] dm(x_2) \right] dm(x_1) \\ &= 1 \end{aligned}$$

(2) We observe that

$$\begin{aligned} [b] &= [b_d | b_{d-1}, \dots, b_1] [b_{d-1} | b_{d-2}, \dots, b_1] \cdots [b_1] \\ &= f_d(b_d) \cdots f_1(b_1) \\ &= f(b) \end{aligned}$$

□

**Exercise 2.2.3.** Set  $A = B^d(0, 1) \cap [0, 1]^d$  (the first orthant of the unit  $d$ -ball) and  $a = 0$ . Then  $A_1 = (0, 1)$ . Choose  $f_1 = 1_{(0,1)}$  and sample  $b_1 \sim f_1$ . For each  $j \in \{2, \dots, d\}$ , set

$$s_j = \sqrt{1 - \sum_{k=1}^{j-1} b_k^2}$$

Then for each  $j \in \{2, \dots, d\}$ ,

$$A_j = (0, s_j)$$

*Proof.* Clear. □

**Exercise 2.2.4.** Continuing from the previous problem, for each  $j \in \{2, \dots, d\}$ , choose  $f_j = s_j^{-1} 1_{(0, s_j)}$ . Then  $f = \left( \prod_{j=2}^d s_j^{-1} \right) 1_{\prod_{j=2}^d (0, s_j)}$ .

*Proof.* Clear. □

**Definition 2.2.5.** Now make  $f_j \sim GP(\mu_j, c_j)$ .

### 2.3. Importance Sampling.

## 2.4. Rejection Sampling.

**Exercise 2.4.1.** Let  $f, g \in \mathcal{D}(\mathbb{R}^d)$  and  $A \in \mathcal{B}(\mathbb{R}^d)$ . Suppose that  $m^d(A) > 0$ . If  $X \sim f$ , then  $X|X \in A \sim \|fI_A\|_1^{-1} fI_A$ .

*Proof.* Let  $C \in \mathcal{B}(\mathbb{R}^d)$ . Then

$$\begin{aligned} P(X \in C|X \in A) &= P(X \in C \cap A)P(X \in A)^{-1} \\ &= \|fI_A\|_1^{-1} \int_C fI_A dm^d \end{aligned}$$

So  $f_{X|X \in A} = \|fI_A\|_1^{-1} fI_A$ . □

**Exercise 2.4.2.** Let  $A, B \in \mathcal{B}(\mathbb{R}^d)$ . Suppose that  $A \subset B$  and  $0 < m^d(A)$  and  $m^d(B) < \infty$ . If  $X \sim \text{Uni}(B)$ , then  $X|X \in A \sim \text{Uni}(A)$ .

*Proof.* Clear using the previous exercise with  $f = I_B$ . □

**Exercise 2.4.3. (Fundamental Theorem of Simulation):**

Let  $f \in \mathcal{D}(\mathbb{R}^d)$  and  $c > 0$ . Define

$$G_c = \{(x, v) \in \mathbb{R}^{d+1} : 0 < v < cf(x)\}$$

- (1) If  $X \sim f$  and  $U \sim \text{Uni}(0, 1)$  are independent, then  $(X, cUf(X)) \sim \text{Uni}(G_c)$ .
- (2) If  $(X, V) \sim \text{Uni}(G_c)$ , then  $X \sim f$ .

*Proof.* First we note that  $m^{d+1}(G_c) = c$ .

- (1) Suppose that  $X \sim f$  and  $U \sim \text{Uni}(0, 1)$  are independent and put  $Y = cUf(X)$ . Then  $Y|X = x \sim cUf(x) \sim \text{Uni}(0, cf(x))$  and we have that for each  $x \in \text{supp } X$  and  $y \in (0, cf(x))$ ,

$$\begin{aligned} f_{X,Y}(x, y) &= f_{Y|X}(y|x)f(x) \\ &= \frac{1}{cf(x)}f(x) \\ &= \frac{1}{c} \end{aligned}$$

So  $(X, Y) \sim \text{Uni}(G_c)$

- (2) Suppose that  $(X, V) \sim \text{Uni}(G_c)$ . Then  $f_{X,V}(x, v) = \frac{1}{c}I_{G_c}(x, v)$ . So

$$\begin{aligned} f_X(x) &= \int_{\mathbb{R}} \frac{1}{c}I_{G_c}(x, v)dm(v) \\ &= \int_0^{cf(x)} \frac{1}{c}dv \\ &= f(x) \end{aligned}$$

So  $X \sim f$ . □

**Exercise 2.4.4.** Let  $f, g \in \mathcal{D}(\mathbb{R}^d)$ ,  $c_f, c_g > 0$  and  $M > 0$ . Put  $\tilde{f} = c_f f$  and  $\tilde{g} = c_g g$ . Suppose that  $\tilde{f} \leq M\tilde{g}$ . If  $Y \sim g$  and  $U \sim \text{Uni}(0, 1)$  are independent, then  $Y|U \leq \frac{\tilde{f}(Y)}{M\tilde{g}(Y)} \sim f$  and  $P(U \leq \frac{\tilde{f}(Y)}{M\tilde{g}(Y)}) = \frac{c_f}{c_g M}$

*Proof.* Put

$$G_g = \{(y, v) \in \mathbb{R}^{d+1} : 0 < v < M\tilde{g}(y)\}$$

and

$$G_f = \{(y, v) \in \mathbb{R}^{d+1} : 0 < v < \tilde{f}(y)\}$$

Then  $G_f \subset G_g$ ,  $m^d(G_g) = c_g M$  and  $m^d(G_f) = c_f$ . By the first part of the fundamental theorem of simulation, we know that

$$(Y, MU_{c_g g}(Y)) \sim \text{Uni}(G_g)$$

Since  $\{(Y, MU_{c_g g}(Y)) \in G_f\} = \{U \leq \frac{c_f f(Y)}{M c_g g(Y)}\}$ , a previous exercise tells us that

$$(Y, MU_{c_g g}(Y))|U \leq \frac{c_f f(Y)}{M c_g g(Y)} \sim \text{Uni}(G_f)$$

Then the second part of the fundamental theorem of simulation tells us that

$$Y|U \leq \frac{c_f f(Y)}{M c_g g(Y)} \sim f$$

Finally we have that

$$\begin{aligned} P\left(U \leq \frac{c_f f(Y)}{M c_g g(Y)}\right) &= P[(Y, MU_{c_g g}(Y)) \in G_f] \\ &= \frac{c_f}{c_g M} \end{aligned}$$

□

**Definition 2.4.5. (Rejection Sampling Algorithm):**

Let  $f, g \in \mathcal{D}(\mathbb{R}^d)$ ,  $c_f, c_g > 0$  and  $M > 0$ . Put  $\tilde{f} = c_f f$  and  $\tilde{g} = c_g g$ . Suppose that  $\tilde{f} \leq M\tilde{g}$ . We define the **rejection sampling algorithm** as follows:

- (1) sample  $Y \sim g$  and  $U \sim \text{Uni}(0, 1)$  independently
- (2) if  $U \leq \frac{\tilde{f}(Y)}{M\tilde{g}(Y)}$ , accept  $Y$ , else return to (1).

If we sample  $(X_n)_{n \in \mathbb{N}}$  independently using the rejection sampler, then the previous exercises imply that  $(X_n)_{n \in \mathbb{N}} \stackrel{iid}{\sim} f$  and the acceptance rate is  $\frac{c_f}{c_g M}$ .

**Note 2.4.6.** Phrasing the rejection sampler in terms of  $\tilde{f}$  and  $\tilde{g}$  instead of  $f$  and  $g$  is useful because we may not always be able to solve for the normalizing constants.

### 3. POSTERIOR CONSISTENCY

#### 3.1. Introduction.

**Definition 3.1.1.** Let  $(\mathcal{X}, \mathcal{F})$  and  $\Theta$  be