

## Exercise – Understanding the theory of rejection sampling

### 1. Aim of the exercise

The aim of this exercise is to develop intuition and understanding of the rejection sampling method. Specifically, we focus on its theoretical foundations, the acceptance criterion, and the role of the proposal distribution and envelope constant in determining efficiency.

### 2. Theory

Our goal is to generate random samples from a specific probability distribution. We focus on the univariate case and represent the distribution by its density function  $f_X(x)$ . This distribution is the target of our sampling procedure. Two widely used methods for sampling from such a density are the inverse transformation method and the rejection sampling method.

The inverse transformation method was introduced in an exercise dedicated to this method. This approach relies on setting the CDF  $F_X(x)$  equal to a uniform random variable  $U$ , and then solving for  $x$  in terms of  $U$ . Specifically, when  $U \sim \text{Unif}(0, 1)$ , the transformation

$$X = F_X^{-1}(U)$$

allows us to generate samples from  $F_X$  by applying the inverse CDF to  $U$ . If the normalized CDF  $F_X$  is known and admits a (generalized) inverse, the inverse transformation method is the most efficient choice. In many practical situations, however, this condition is not satisfied. A more flexible technique with less restrictive assumptions is rejection sampling, also known as the accept-reject method.

The inverse transformation method belongs to the class of direct sampling techniques, since it operates directly on the CDF of the random variable  $X$  to be generated. Rejection sampling, introduced by von Neumann (1951), is an indirect sampling technique. Unlike the inverse transformation method, rejection sampling does not require the target density to be normalized, that is, to integrate to 1. In many applications we can evaluate a function  $f_X(x)$  that is proportional to the target density. It has the correct shape but is not itself a probability density, because we do not know the normalizing constant that makes it integrate to 1. Formally, the normalized target density is

$$\tilde{f}_X(x) = \frac{1}{c_f} f_X(x),$$

where  $c_f$ , which is typically unknown, is the normalization constant. Rejection sampling can work directly with the unnormalized function  $f_X$ . It never needs the density to integrate to 1, because the rejection rule uses only ratios of density values where any unknown constant cancels out automatically.

Rejection sampling works by generating trial values from a proposal density  $p_X(x)$ , from which sampling and evaluation are straightforward. The proposal must satisfy

$$f_X(x) \leq C p_X(x),$$

for all  $x \in \mathbb{R}$ , for some finite constant  $C \geq 1$ , so that  $C p_X(x)$  forms an envelope for the target density  $f_X(x)$ . In geometric terms, acceptance amounts to keeping a trial value only when it lies under the graph of the target density relative to this envelope. A smaller  $C$  makes the envelope curve shorter and tighter, so more points fall under the target and are therefore accepted.

Formally, the acceptance step proceeds in three stages. First, draw a trial value  $X^*$  from the proposal density  $p_X$ . This is the candidate point we may accept or reject.

Second, compute the ratio

$$r(x) = \frac{f_X(x)}{C p_X(x)}.$$

Because  $f_X$  and  $p_X$  are nonnegative densities and the envelope condition  $f_X(x) \leq C p_X(x)$  holds, dividing by  $C p_X(x)$  ensures that

$$0 \leq r(x) \leq 1.$$

This ratio is therefore a valid probability, interpreted as the chance of accepting a candidate located at  $x$ .

Third, we need a mechanism to turn this probability into a concrete accept or reject decision. For this purpose, draw an independent uniform random variable  $U \sim \text{Unif}(0, 1)$ . A uniform variable provides exactly such a mechanism, because the event  $\{r \geq U\}$  occurs with probability  $r$  for any  $r \in [0, 1]$ . Therefore the trial value  $X^*$  is accepted whenever

$$U \leq \frac{f_X(X^*)}{C p_X(X^*)},$$

and rejected otherwise in accordance with the requirement that the proposal density must satisfy the envelope inequality.

The correctness of rejection sampling can be seen by looking at the joint distribution of the candidate  $X^*$  and the uniform variable  $U$ . Since  $X^*$  is drawn from the proposal density  $p_X(x)$  and  $U$  is drawn independently from  $\text{Unif}(0, 1)$ , their joint density is

$$f_{X^*, U}(x, u) = p_X(x) 1_{[0,1]}(u).$$

A candidate  $x$  is accepted if

$$U \leq \frac{f_X(x)}{C p_X(x)}.$$

Therefore, for each  $x$ , the accepted values of  $u$  lie in the interval

$$0 \leq u \leq \frac{f_X(x)}{C p_X(x)}.$$

Next, consider the probability of acceptance in a set  $A$ . Let  $\mathcal{A}$  denote the event that a proposal is accepted. The probability that the proposed value lies in a measurable set  $A \subseteq \mathbb{R}$  and  $\mathcal{A}$  occurs is

$$\Pr((X^* \in A) \cap \mathcal{A}) = \int_A \int_0^{\frac{f_X(x)}{C p_X(x)}} p_X(x) du dx.$$

Evaluating the inner integral gives

$$\int_0^{\frac{f_X(x)}{C p_X(x)}} du = \frac{f_X(x)}{C p_X(x)}.$$

Simplifying, we obtain

$$\Pr((X^* \in A) \cap \mathcal{A}) = \frac{1}{C} \int_A f_X(x) dx.$$

The unconditional acceptance probability is

$$\Pr(\mathcal{A}) = \int_{\mathbb{R}} p_X(x) \frac{f_X(x)}{C p_X(x)} dx = \frac{1}{C}.$$

Conditioning on acceptance, we obtain

$$\Pr(X^* \in A | \mathcal{A}) = \frac{\Pr((X^* \in A) \cap \mathcal{A})}{\Pr(\mathcal{A})} = \int_A f_X(x) dx.$$

This shows that the accepted draws have density  $f_X$ . The acceptance probability is  $\frac{1}{C}$ , which highlights the importance of choosing a proposal density  $p_X$  that approximates the shape of  $f_X$  as closely as possible.

The rejection sampling mechanism can be summarized in the following pseudo-code:

---

**Algorithm:** Pseudo-code for rejection sampling

---

```

1: Input: number of iterations  $n$ , scaling constant  $C$ 
2: for  $i = 1, \dots, n$ 
3:   Generate  $X^* \sim p_X$ 
4:   Generate  $U \sim \text{Unif}(0, 1)$ 
5:   if  $U \leq \frac{\tilde{f}_X(X^*)}{C \cdot p_X(X^*)}$  then
6:     Accept  $X^*$ 
7:   else Reject  $X^*$ 
8:   end if
9: end for
```

---

This acceptance-rejection step is conceptually related to the Metropolis–Hastings algorithm, since both employ a probabilistic acceptance criterion. The crucial difference is that rejection sampling discards rejected proposals, whereas Metropolis–Hastings retains them as part of the chain. As a result, rejection sampling can be less efficient computationally. A more detailed discussion of Metropolis–Hastings is provided in a dedicated exercise.

The following theorem establishes the validity of the rejection sampling method.

**Theorem 1.** *Let  $X$  be the random variable generated by the rejection sampling algorithm. Then  $X$  has the desired target probability density function  $f_X$ .*

*Proof.* See Rubinstein (2016). ■

Theorem 1 applies to any proposal density  $p_X$  that satisfies the following conditions:

- 
- RS.1 Support condition:  $\text{Supp}(\tilde{f}_X) \subseteq \text{Supp}(p_X)$ .
  - RS.2 Envelope condition: there exists a constant  $C \in \mathbb{R}_{\geq 1}$ .  
such that  $\tilde{f}_X(x) \leq C \cdot p_X(x)$  for all  $x \in \mathbb{R}$ .
- 

### 3. Final notes

This file is prepared and copyrighted by Jelmer Wieringa and Tunga Kantarci. This file is

available on GitHub and can be accessed using this [link](#). We used the following reference. Von Neumann, J. (1951) Various techniques used in connection with random digits. In: Householder, A.S., Forsythe, G.E. and Germond, H.H. (eds.) Monte Carlo Method. National Bureau of Standards Applied Mathematics Series, No. 12. Washington, DC: U.S. Government Printing Office, pp. 36-38.