PROPOSAL DISTRIBUTION

Machine Learning

IIIy CSE IDP and IDDMP

Refer : Stephen Marsland

Motivation

1. Direct sampling from p(x) is often computationally expensive or infeasible.
2. We can evaluate an un-normalised version p'(x) but cannot draw samples easily.
3. Introduce a simpler, easy-to-sample “proposal” distribution q(x).
4. **Goal**: use samples from q(x) to approximate samples from p(x) without altering notations.

Proposal Distribution

1. We assume

$$p(\mathbf{x}) = \frac{1}{Z_p}\tilde{p}(\mathbf{x})$$

with unknown normalisation constant Zp

2. Choose q(x) such that for all x. $\tilde{p}(x) \ \leq \ M\, q(x)$

where M is a finite constant (the “envelope” factor).

3. Sampling procedure: draw $x^{*} \sim q(x)$, then decide acceptance via a uniform test under $M\, q(x^{*})$

The Rejection Sampling Algorithm

The Rejection Sampling Algorithm

- Sample $\mathbf{x}^*$ from $q(\mathbf{x})$ (e.g., using the Box–Muller scheme if $q(\mathbf{x})$ is Gaussian)
- Sample u from uniform$(0, \mathbf{x}^*)$
- If $u < p(\mathbf{x}^*)/Mq(\mathbf{x}^*)$:
  - add $\mathbf{x}^*$ to the set of samples
- Else:
  - reject $\mathbf{x}$ and pick another sample

Rejection Sampling Example & Transition to Importance Sampling

Example (Mixture of Gaussians): (Refer next slide for graph)

Proposal q(x) uniform (dotted line).

With M=0.8, ≈50% of samples are rejected. With M=2, ≈85% of samples are rejected.

Issues:

A) High rejection rates waste computation. B)Curse of dimensionality exacerbates inefficiency.

Two Remedies:

Develop more sophisticated exploration methods. Bias sampling toward high-probability regions.

Importance Sampling:

Assign each draw $x^{(i)} \sim q(x)$ a weight $$w\left(x^{(i)}\right) = \frac{p\left(x^{(i)}\right)}{q\left(x^{(i)}\right)}$$

Allows estimation of expectations without discarding samples.

Resampling:

Use weights in a **Sampling-Importance-Resampling** scheme to obtain a representative set of points.

$$E(f) \approx \frac{1}{N} \sum_{i=1}^{N} \frac{p\left(\mathbf{x}^{(i)}\right)}{q\left(\mathbf{x}^{(i)}\right)} f\left(x^{(i)}\right)$$

Rejection Sampling Example & Transition to Importance Sampling

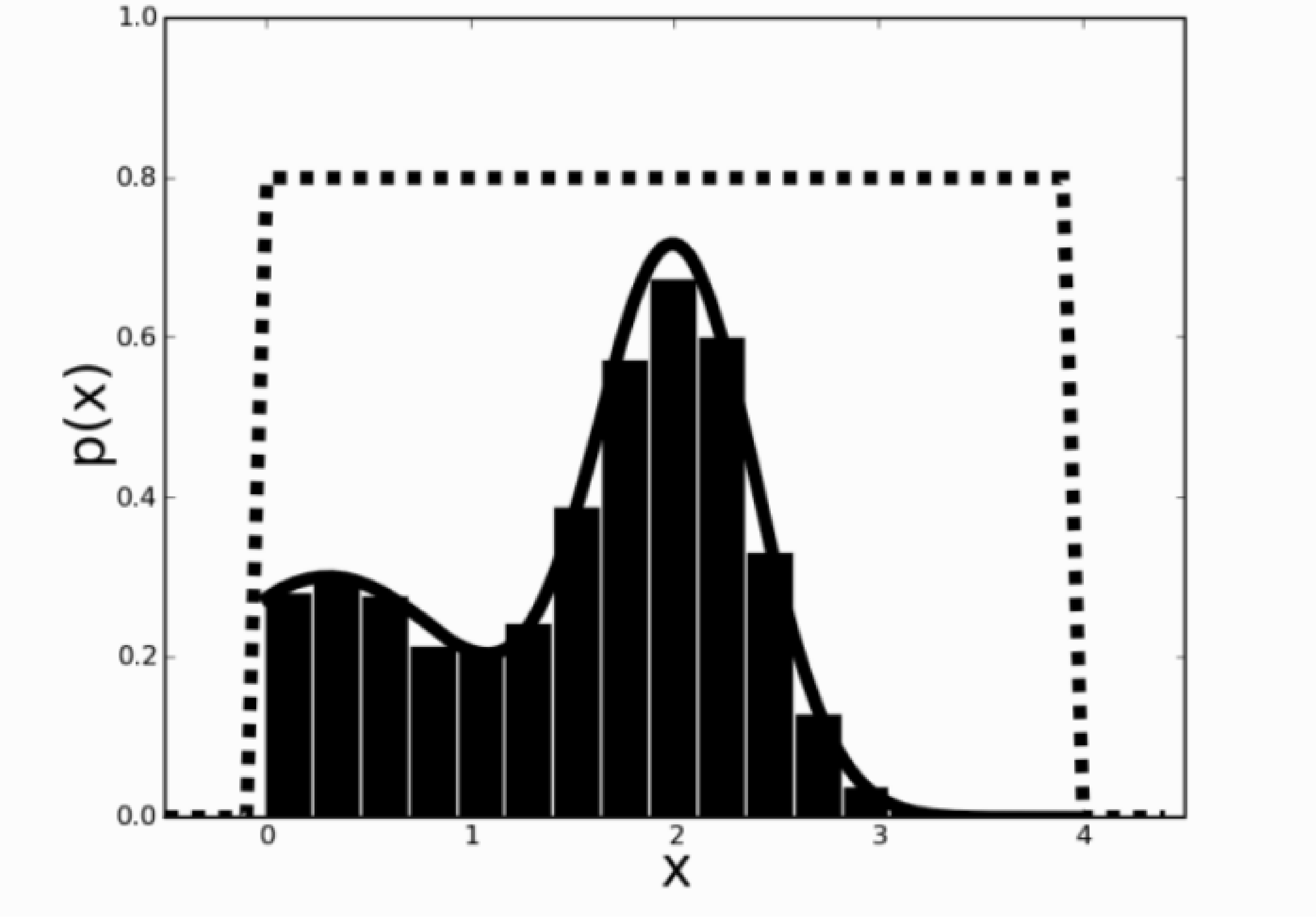

FIGURE 15.3 The histogram shows samples of a mixture of two Gaussians (given by the solid line) as sampled from the uniform box shown as a dotted line by using rejection sampling.

The Sampling-Importance-Resampling Algorithm

The Sampling-Importance-Resampling Algorithm

- Produce N samples $\mathbf{x}^{(i)}, \; i = 1 \ldots N$ from $q(\mathbf{x})$
- Compute normalised importance weights

$$w^{(i)} = \frac{p(\mathbf{x}^{(i)})/q(\mathbf{x}^{(i)})}{\sum_j p(\mathbf{x}^{(j)})/q(\mathbf{x}^{(j)})}$$

- Resample from the distribution $\{\mathbf{x}^{(i)}\}$ with probabilities given by the weights $w^{(i)}$

Summary And Conclusion

Challenge: Direct sampling from complex **p(x)** is often infeasible.

Rejection Sampling: Uses a simpler proposal **q(x)** and an envelope **M q(x)** to filter samples.
Easy to implement but can discard many draws, especially in high dimensions.

Improvements:

Better exploration strategies (local moves, MCMC).

Importance sampling to weight rather than reject.

Importance Sampling:

Assigns importance weights to all samples, avoiding waste.

Enables weighted estimation and resampling (SIR).

Takeaway:

Choice of **q(x)** and **M** is critical for efficiency.

Importance sampling offers a powerful alternative when outright rejection is too costly