Lecture 9: Monte Carlo estimates of various statistics

Professor Ilias Bilionis

Estimating the probability density function via histograms

PREDICTIVE
SCIENCE LABORATORY

Estimating the probability density function via histograms

- Let X be a random variable and $Y = g(X)$.
- We want to approximate the PDF $p(y)$ of Y.

PREDICTIVE
SCIENCE LABORATORY

Estimating the probability density function via histograms

- Take M small bins $[b_0, b_1], \dots, [b_{M-1}, b_M]$ in the y space.
- We will approximate $p(y)$ with a constant inside each bin:

$$\hat{p}_M(y) = \sum_{j=1}^{M} c_j 1_{[b_{j-1}, b_j]}(y),$$

- Where the c_j's are coefficients to be determined.

PREDICTIVE SCIENCE LABORATORY

Estimating the probability density function via histograms

- We will approximate $p(y)$ with a constant inside each bin:

$$\hat{p}_M(y) = \sum_{j=1}^{M} c_j 1_{[b_{j-1}, b_j]}(y),$$

- The constants are just:

$$c_j = p(b_{j-1} \leq Y \leq b_j) = F(b_j) - F(b_{j-1})$$

$$\approx \bar{F}_N(b_j) - \bar{F}_N(b_{j-1})$$

PREDICTIVE
SCIENCE LABORATORY

Estimating the probability density function via histograms

- So, we can approximate the constants with the empirical CDF:

$$\bar{c}_{j,N} := \bar{F}_N(b_j) - \bar{F}_N(b_{j-1}) \rightarrow c_j \text{ a.s.}$$

- The intuitive meaning of this is:

$$\bar{c}_{j,N} = \frac{\text{number of samples that fall in bin } [b_{j-1}, b_j]}{N}$$

PREDICTIVE
SCIENCE LABORATORY

Estimating the probability density function via histograms

- The intuitive meaning of this is:

$$\bar{c}_{j,N} = \frac{\text{number of samples that fall in bin } [b_{j-1}, b_j]}{N}$$

- Putting everything together our approximation becomes:

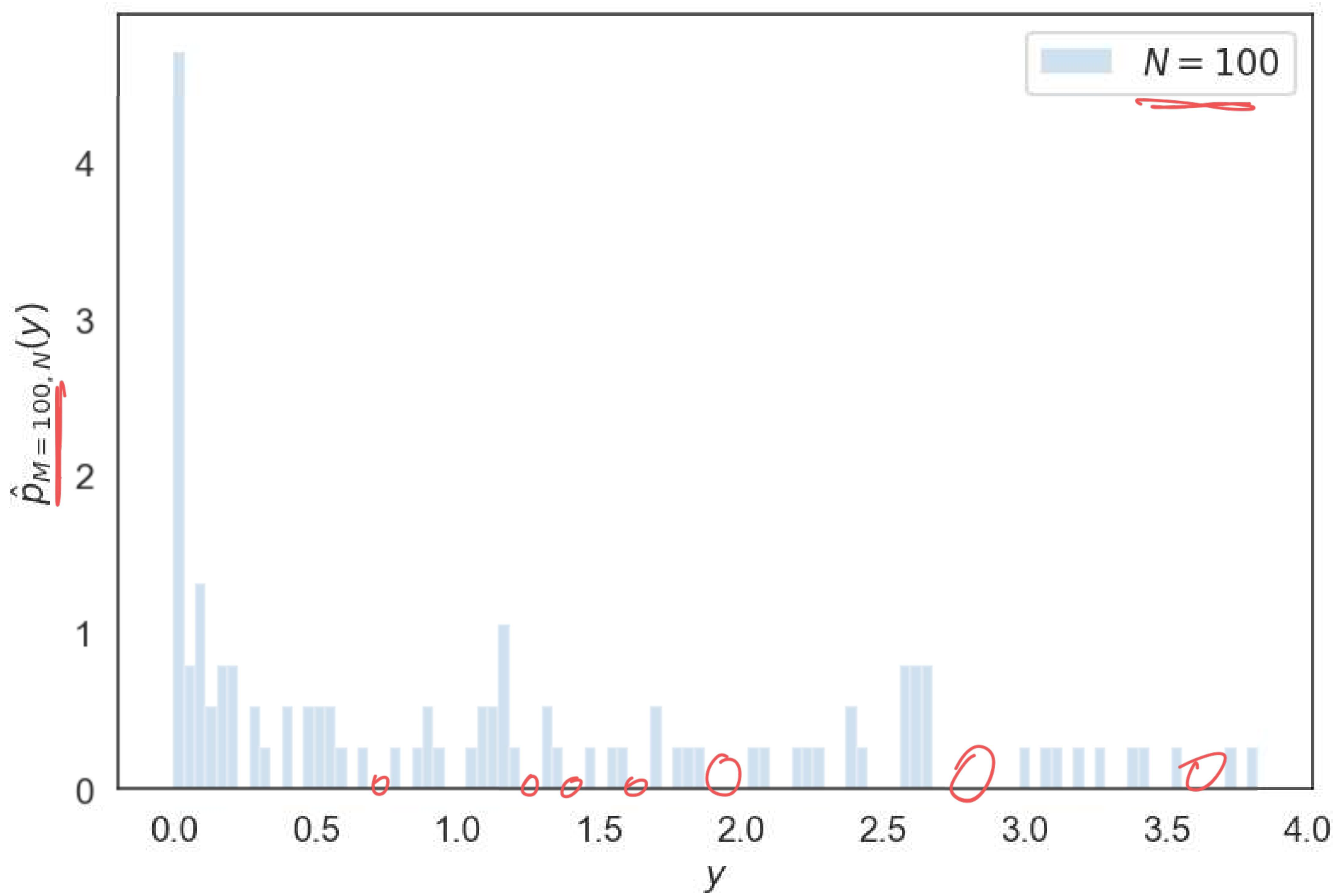
$$\boxed{\hat{p}_{M,N}(y) = \sum_{j=1}^{M} \bar{c}_{j,N} 1_{[b_{j-1},b_j]}(y)}$$

bins

samples

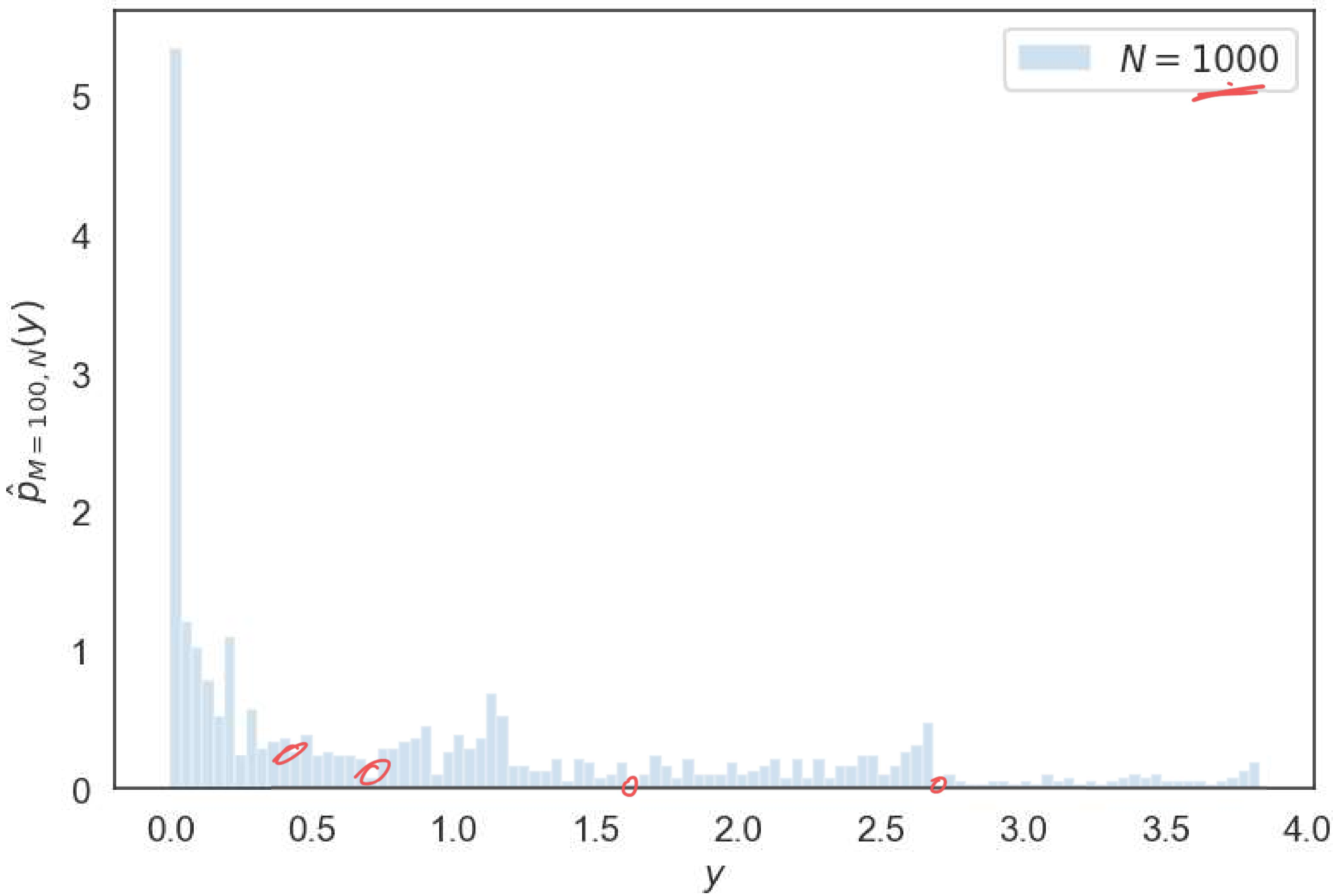
PREDICTIVE
SCIENCE LABORATORY

Example: 1D PDF



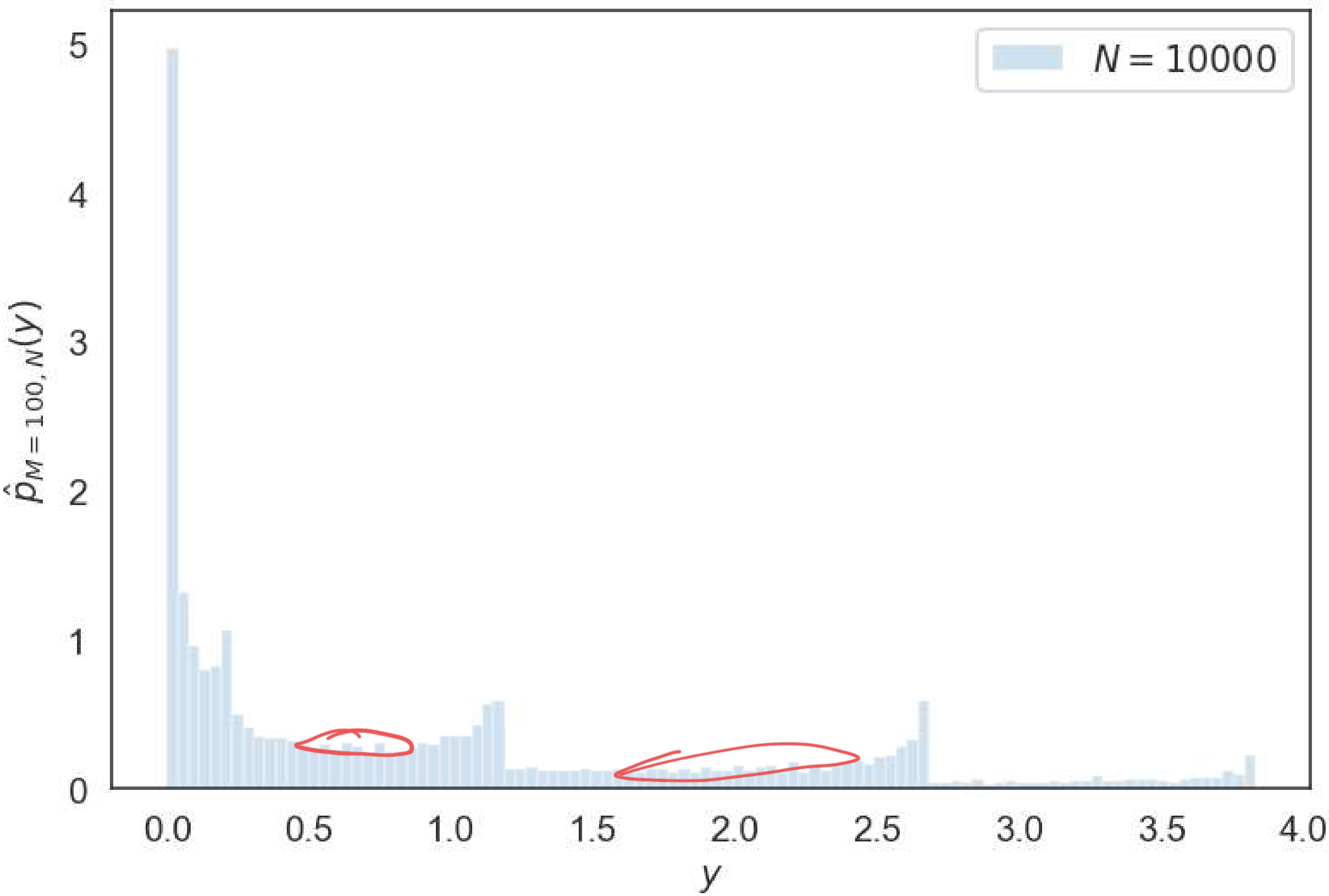
PREDICT
SCIENCE LABORATORY

Example: 1D PDF



PREDICT
SCIENCE LABORATORY

Example: 1D PDF



PREDICT
SCIENCE LABORATORY