



European
Commission



Joint Research Centre

the European Commission's
in-house science service

Joint
Research
Centre

Eighth International Conference on Sensitivity Analysis of Model Output

University of La Réunion Island

30 November, 2016

Sensitivity analysis under different
distributions using the same simulation

Ricardo Bolado-Lavin
Stefano Tarantola

Directorate for Energy, Transport and
Climate - Joint Research Centre of the
European Commission



When experts propose different input distributions, or an a-priori input distribution is updated ...

... we want to see whether changes in the input distribution have an effect on the output distribution.

... and we may want to repeat the Sensitivity Analysis under the new input distributions



The usual practice is to resample from the new input distribution and re-run the code.

If the cost of running the model is high, this practice may be prohibitive.

Beckman and McKay (1987)

Propose two methods that allow the analyst to change the distribution of the inputs without re-running the computer code.

- The weighting method
- The acceptance/rejection method

1. Beckman R.J., McKay M.D. 'Monte Carlo Estimation Under Different Distributions Using the Same Simulation'. **Technometrics** Vol 29, No 2, pages 153-160. 1987.

\mathbf{X} = *vector of model inputs*

$f_1(\mathbf{X})$ = *reference pdf*

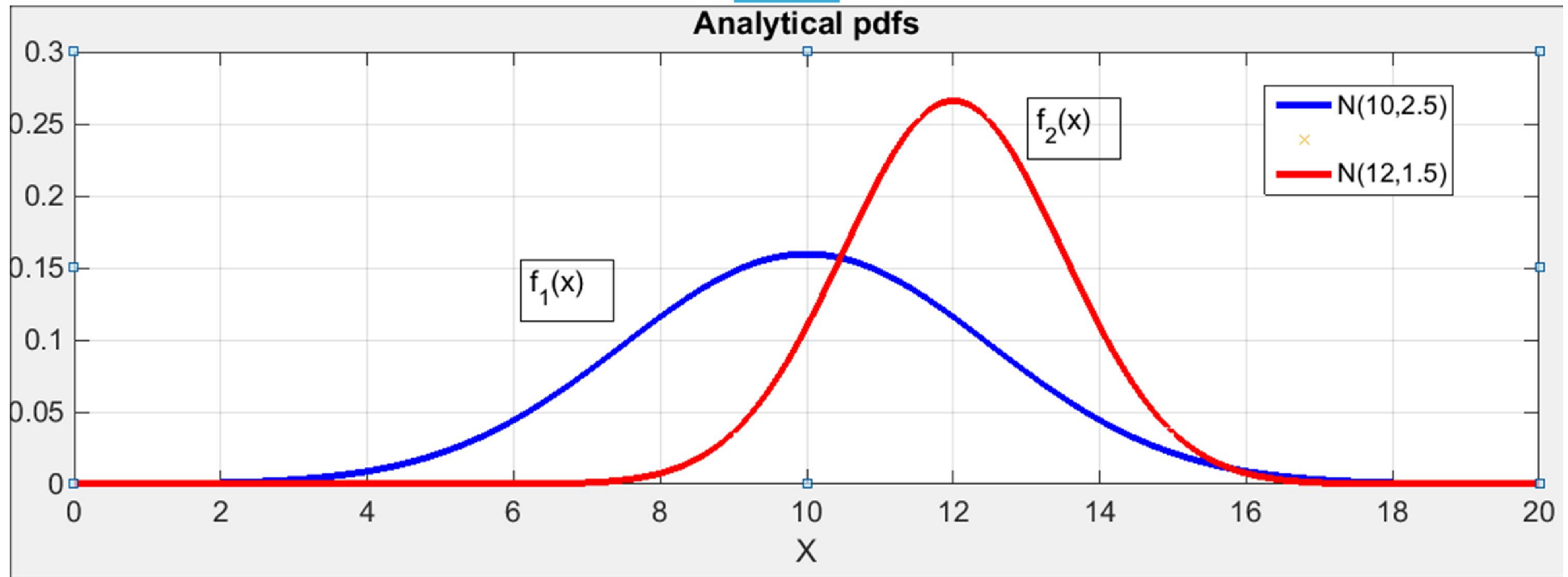
$f_2(\mathbf{X})$ = *alternative pdf*

$Y = H(\mathbf{X})$ = *model output*

The rejection method. Hypotheses:

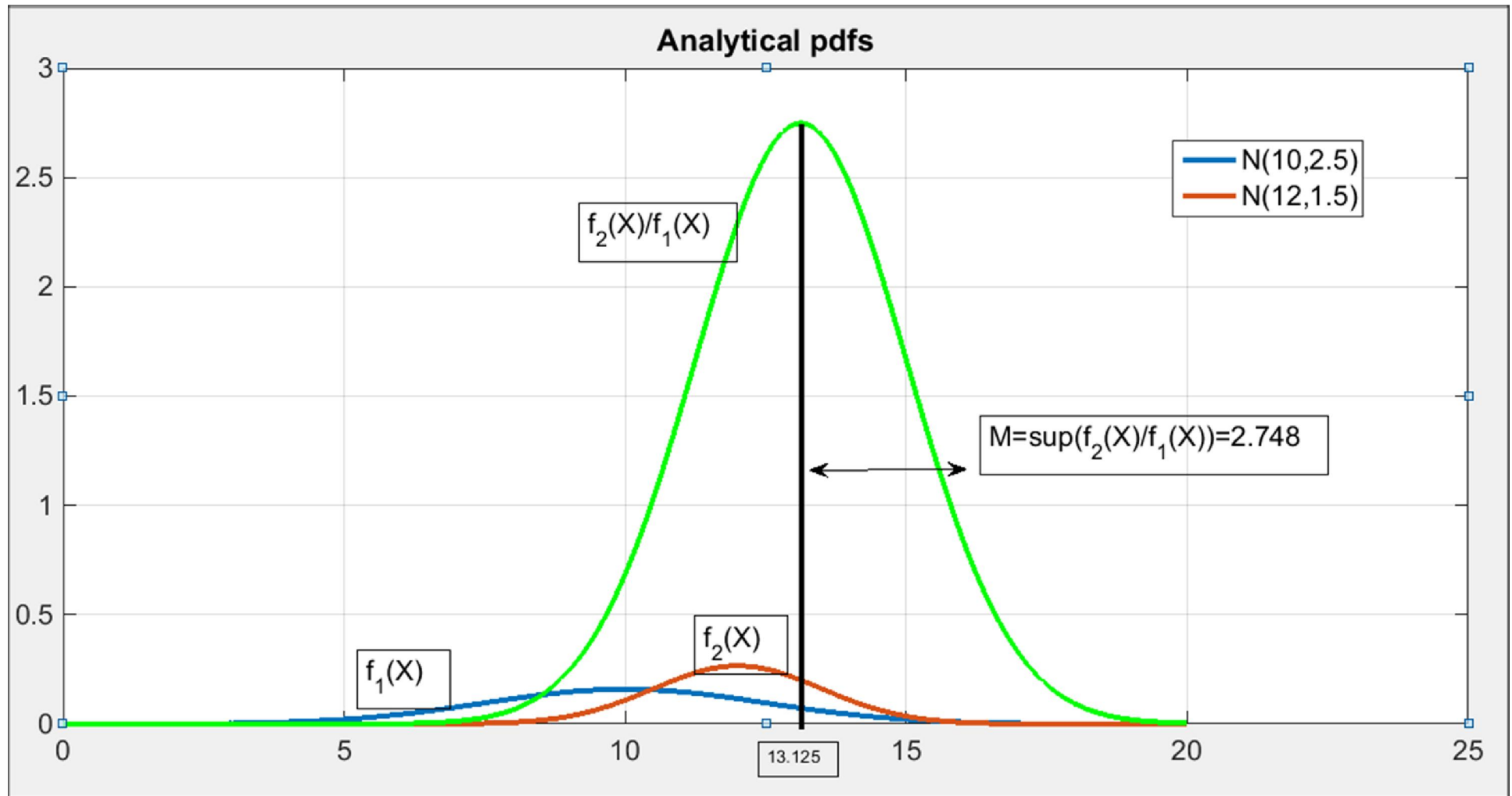
1. $\text{supp}f_2(X)$ is contained in $\text{supp}f_1(X)$
2. $\frac{f_2(X)}{f_1(X)} \leq M$ over $\text{supp}f_2(X)$ with $M < \infty$

Example:



1. $\text{supp} f_2(X)$ is contained in $\text{supp} f_1(X)$

2. $\frac{f_2(X)}{f_1(X)} \leq M$ over $\text{supp} f_2(X)$ with $M < \infty$



Step 1: Generate a sample (x_1, x_2, \dots, x_N) of size N under $f_1(X)$

Step 2: Run the code $Y = H(X)$ on (x_1, x_2, \dots, x_N) and get (y_1, y_2, \dots, y_N)

Step 3: For each sample x_i extract a random number μ_i from $\text{Unif}(0, M f_1(x_i))$

Step 4: If $\mu_i \leq f_2(x_i)$ retain x_i (and y_i) otherwise reject x_i

Step 5: Use the remaining y_i as sample of Y under $f_2(X)$ and build the cdf of Y

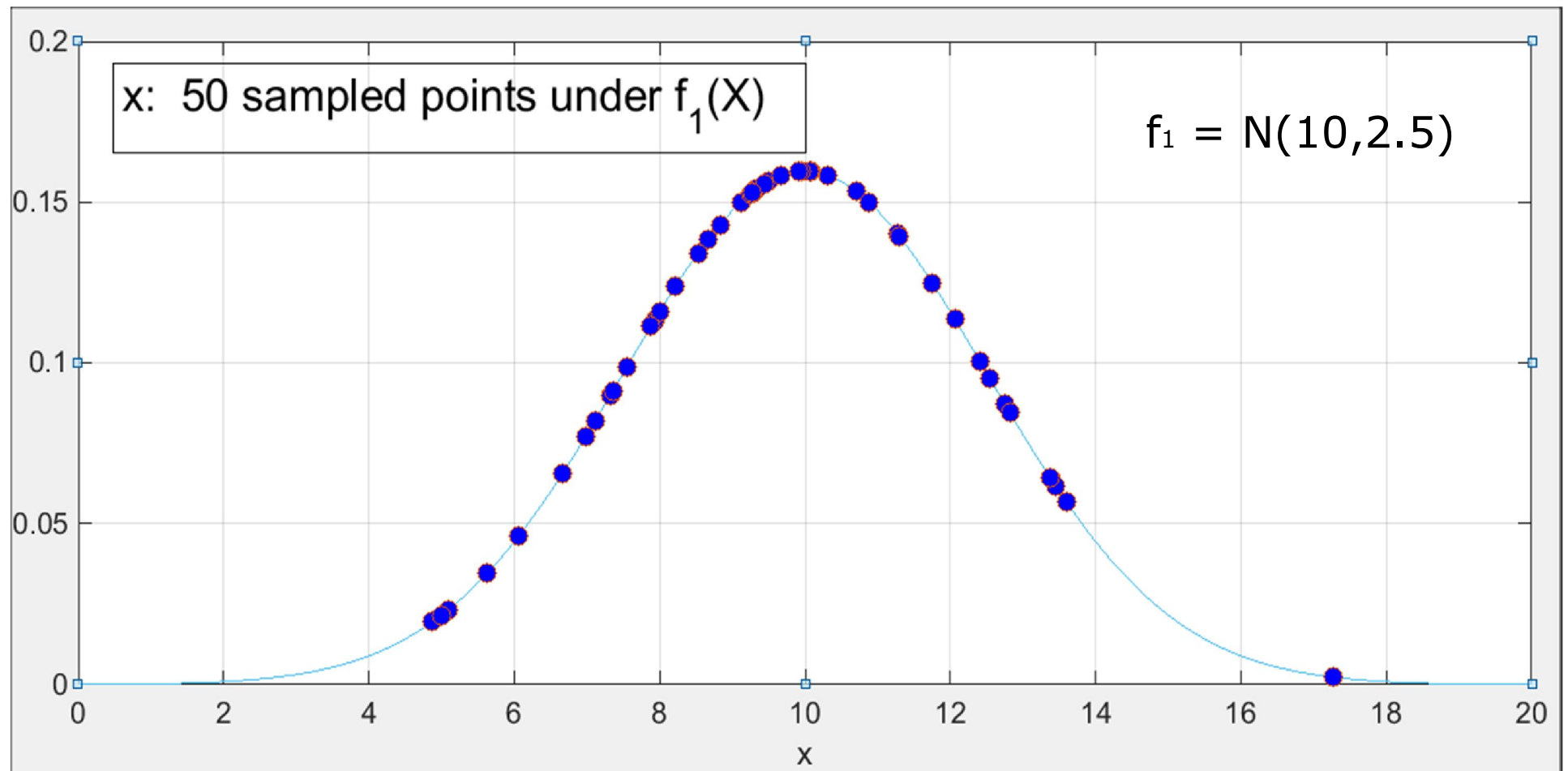
Here, we want to see how the sample (x_1, x_2, \dots, x_N) actually generated under $f_1(X)$ is representative of $f_2(X)$.

Therefore, we skip step 2 and step 5

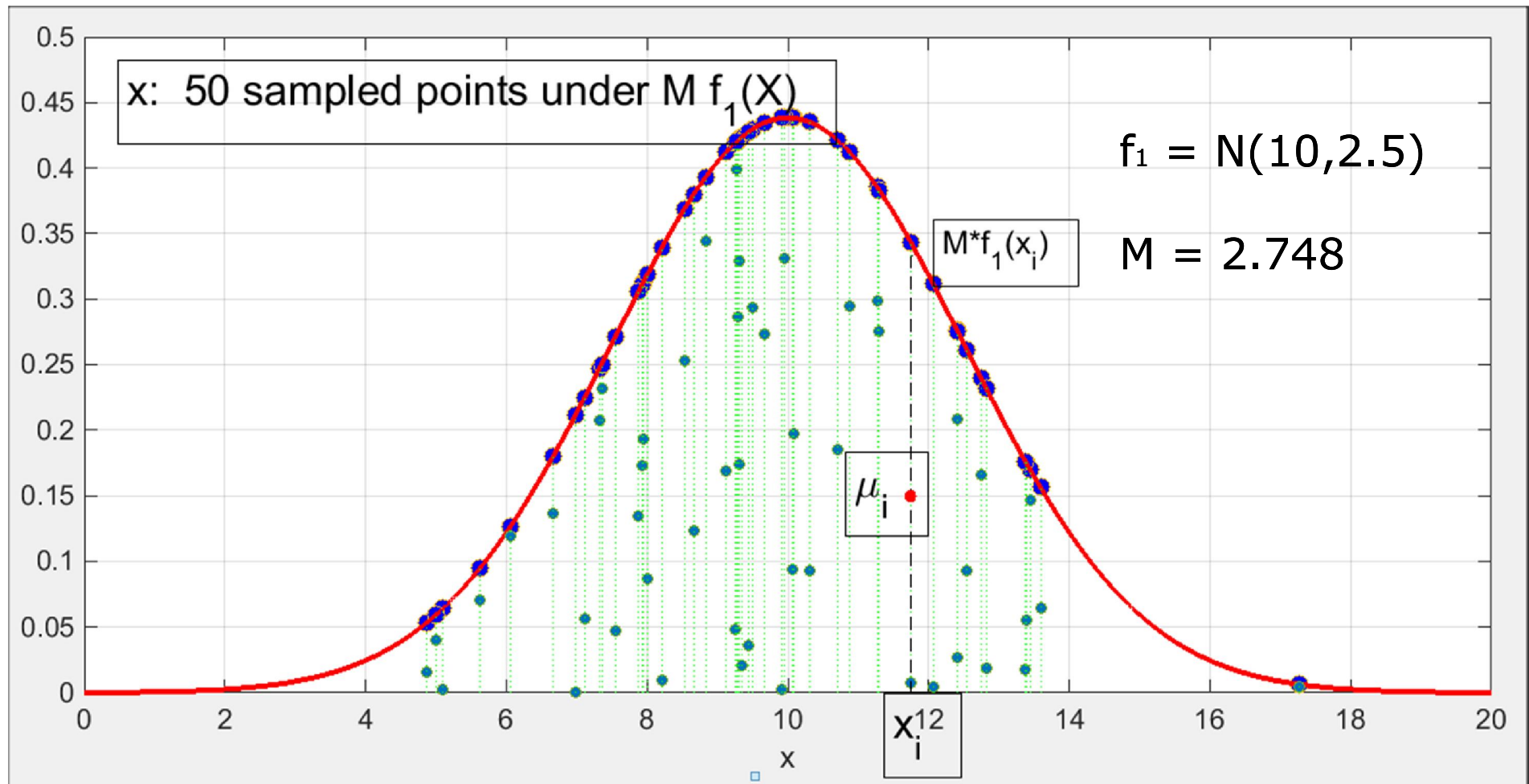
Step 2: Run the code $Y = H(X)$ on (x_1, x_2, \dots, x_N) and get (y_1, y_2, \dots, y_N)

Step 5: Use the remaining y_i as sample of Y under $f_2(X)$ and build the cdf of Y

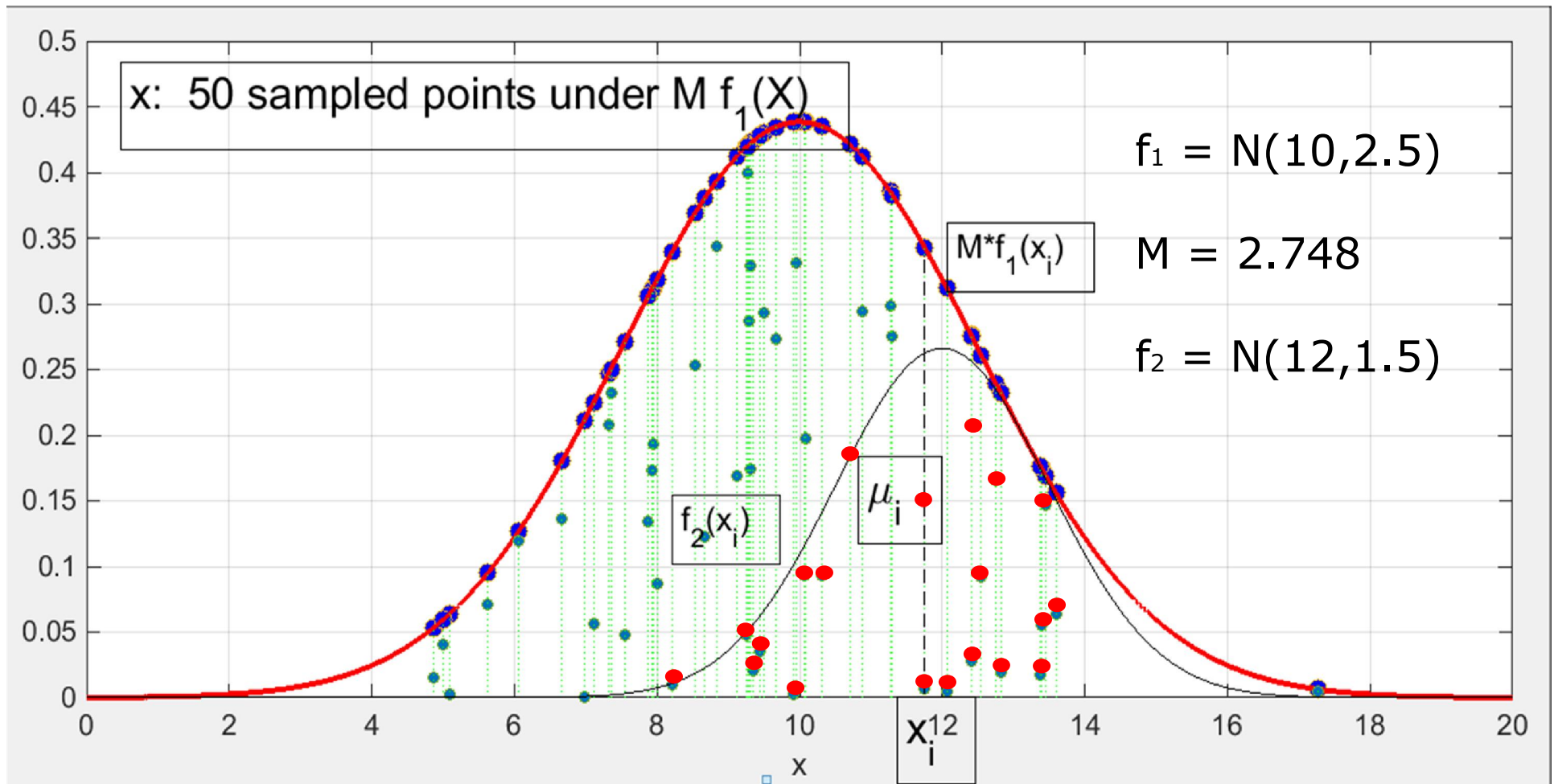
Step 1: Generate a sample (x_1, x_2, \dots, x_N) of size N under $f_1(X)$



Step 3: For each sample x_i extract a random number μ_i from $\text{Unif}(0, M f_1(x_i))$



Step 4: If $\mu_i \leq f_2(x_i)$ retain x_i (and y_i)
otherwise reject x_i



Step 1: Generate a sample (x_1, x_2, \dots, x_N) of size N under $f_1(X)$

Step 2: Run the code n times and get (y_1, y_2, \dots, y_n)

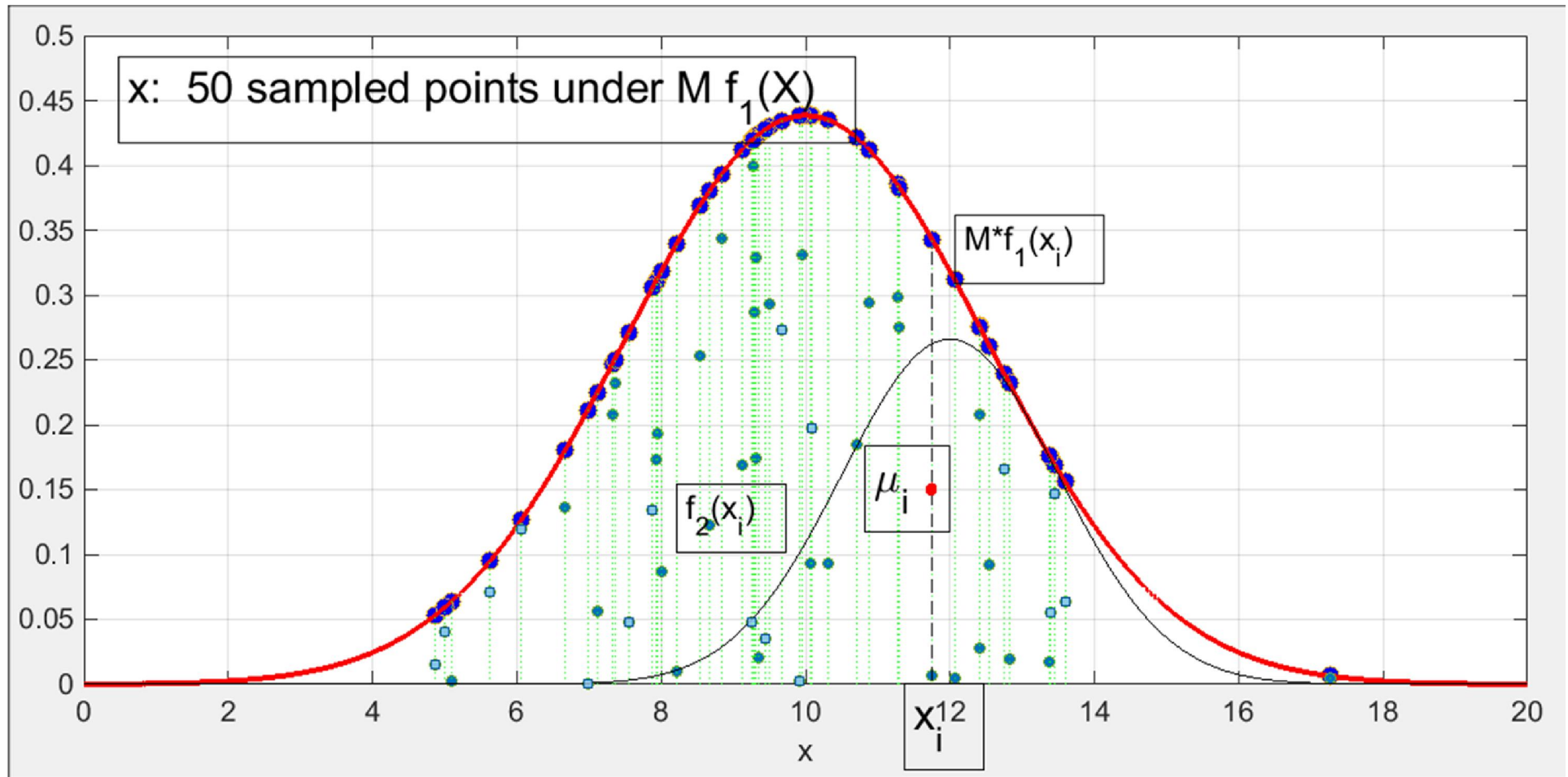
Step 3: For each sample x_i extract a random number μ_i from $\text{Unif}(0, M f_1(X))$

loop

Step 4: If $\mu_i \leq f_2(x_i)$ retain x_i (and y_i) otherwise reject x_i

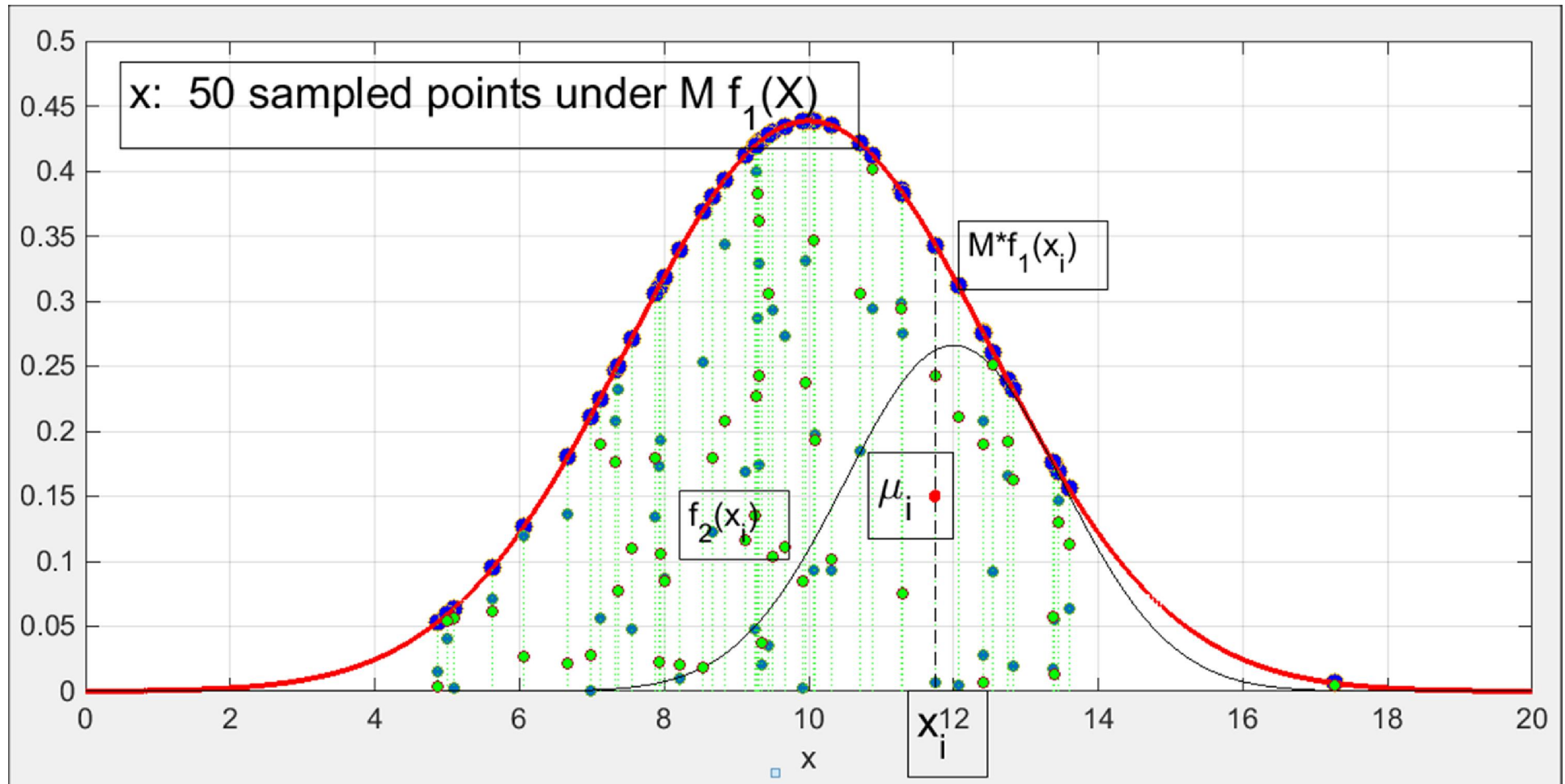
Step 5: Use the remaining y_i as sample of Y under $f_2(X)$ and build the cdf of Y

Loop step3 – step 4 No loop



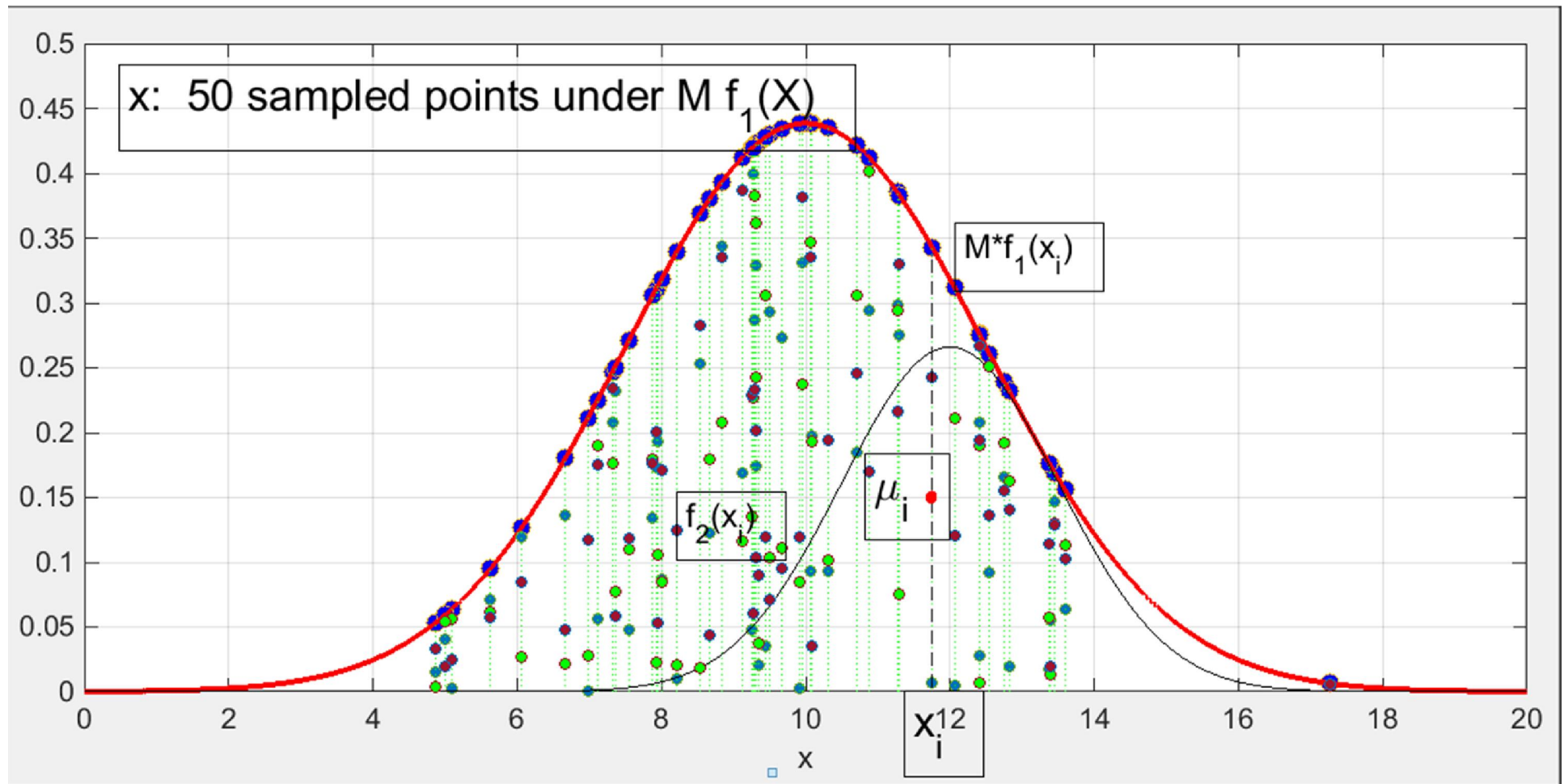
Loop step3 – step 4

One loop



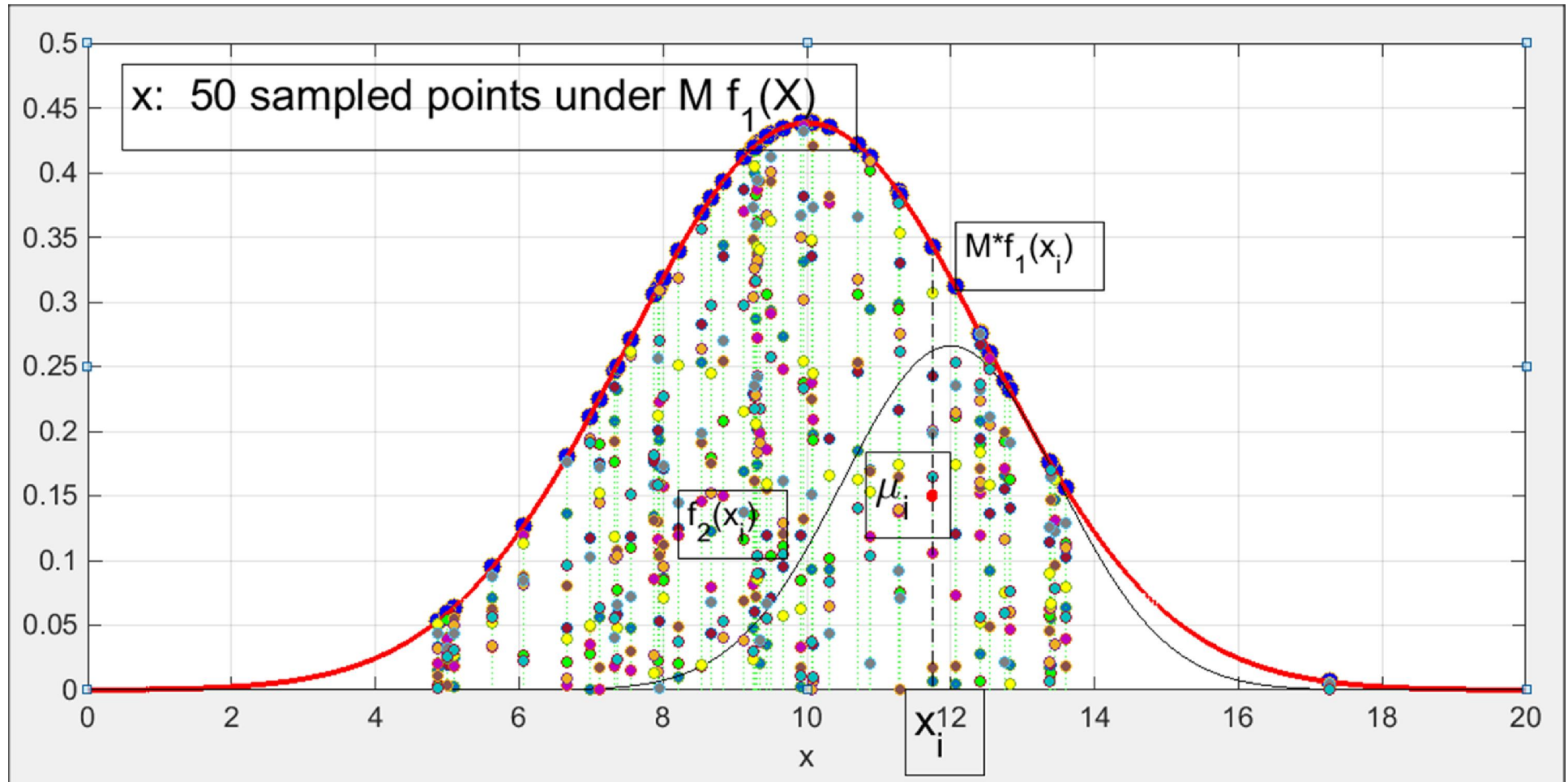
Loop step3 – step 4

Two loops

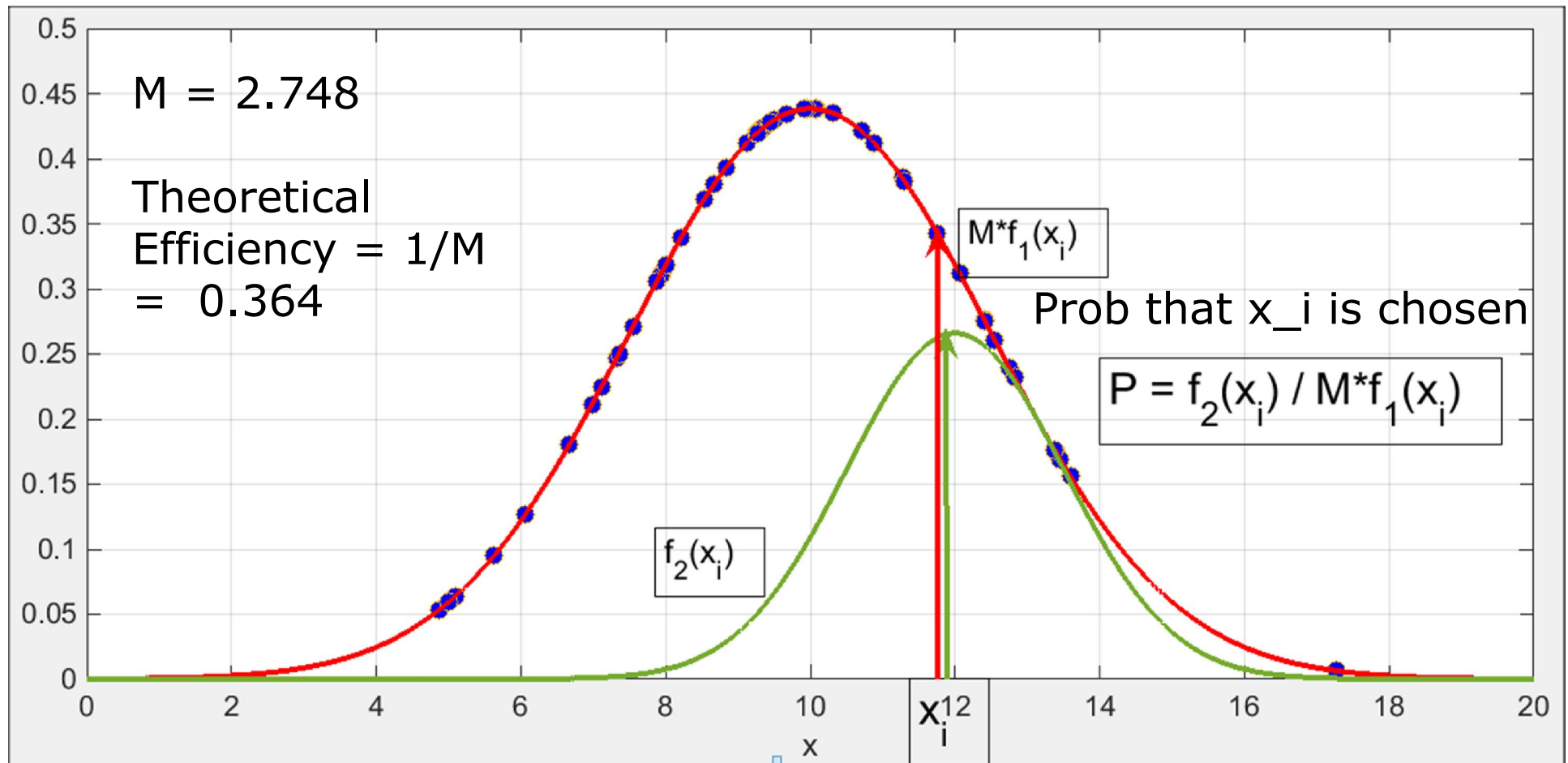


Loop step3 – step 4

Ten loops



Loop step3 – step 4
10, 100, 1000, ... times



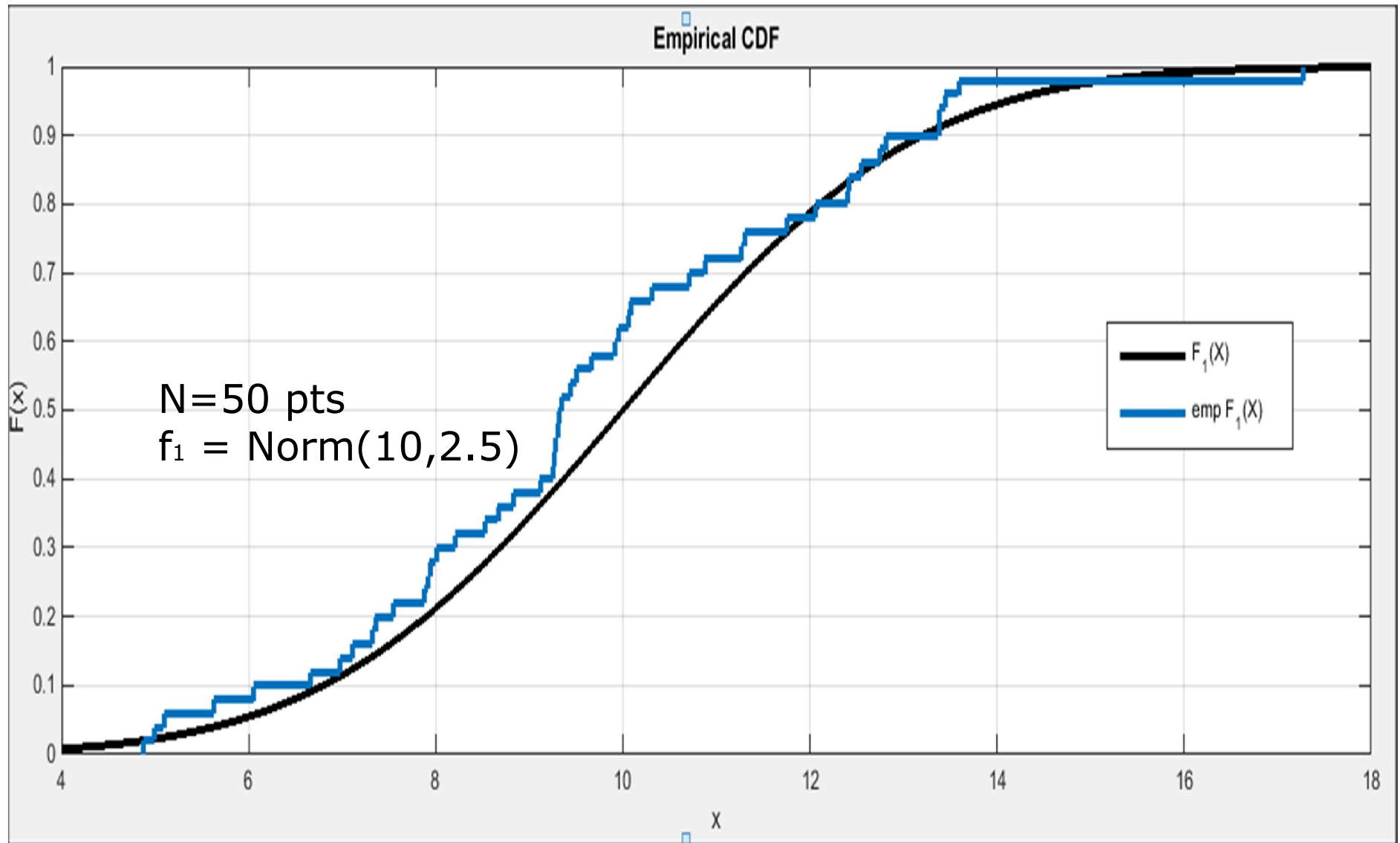


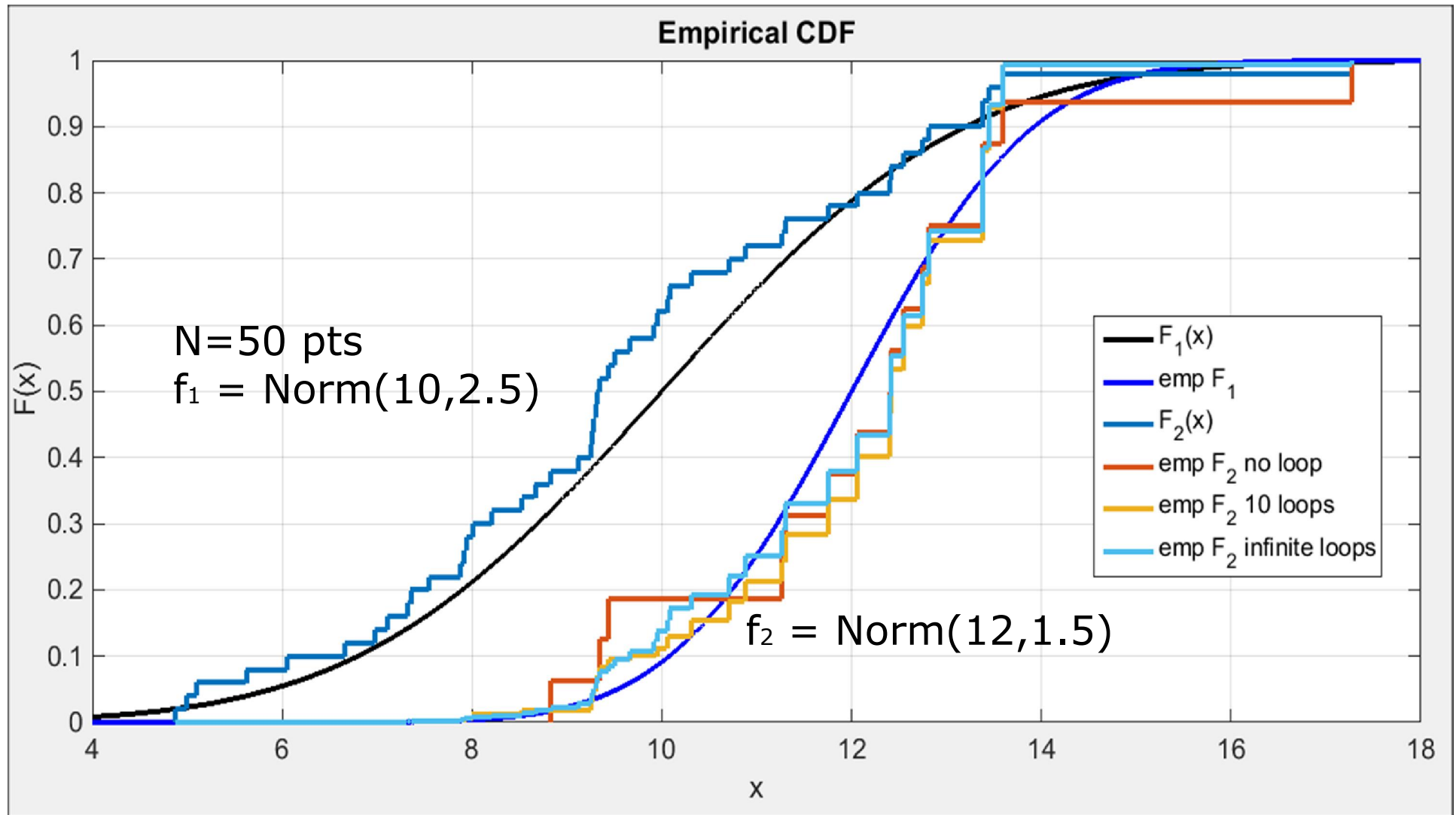
Loop step3 – step 4
10, 100, 1000, ... times

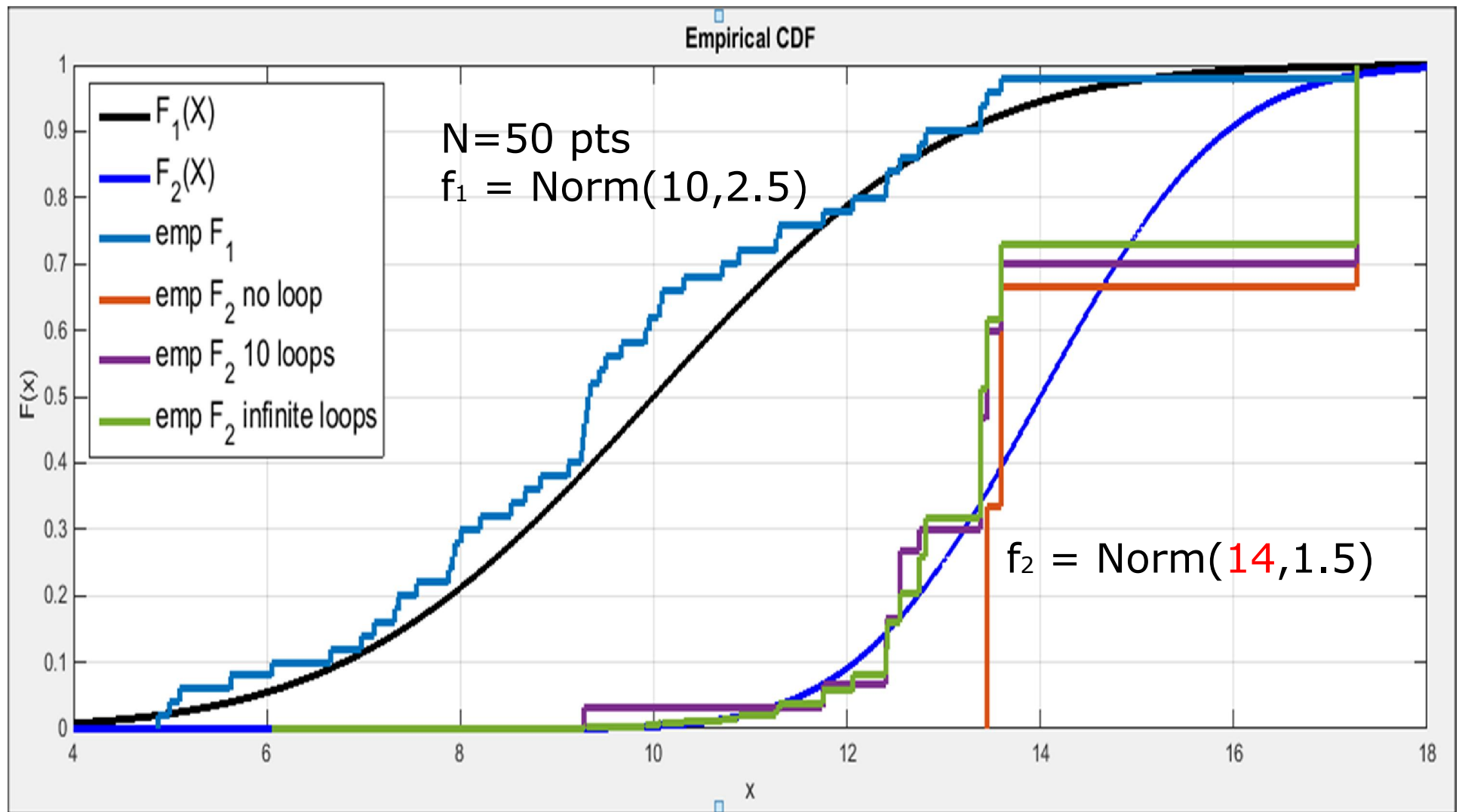
Example:

Without loop: 13 pts remain out of 50
With 10 loops 159 pts remain out of 500
With 100 loops 1,563 pts remain out of 5,000
With 1,000 loops 15,701 pts remain out of 50,000,
etc.

Remember that the real sample is made of only 50
pts. So these are counted many times







Sensitivity Analysis under different distributions using the same sample

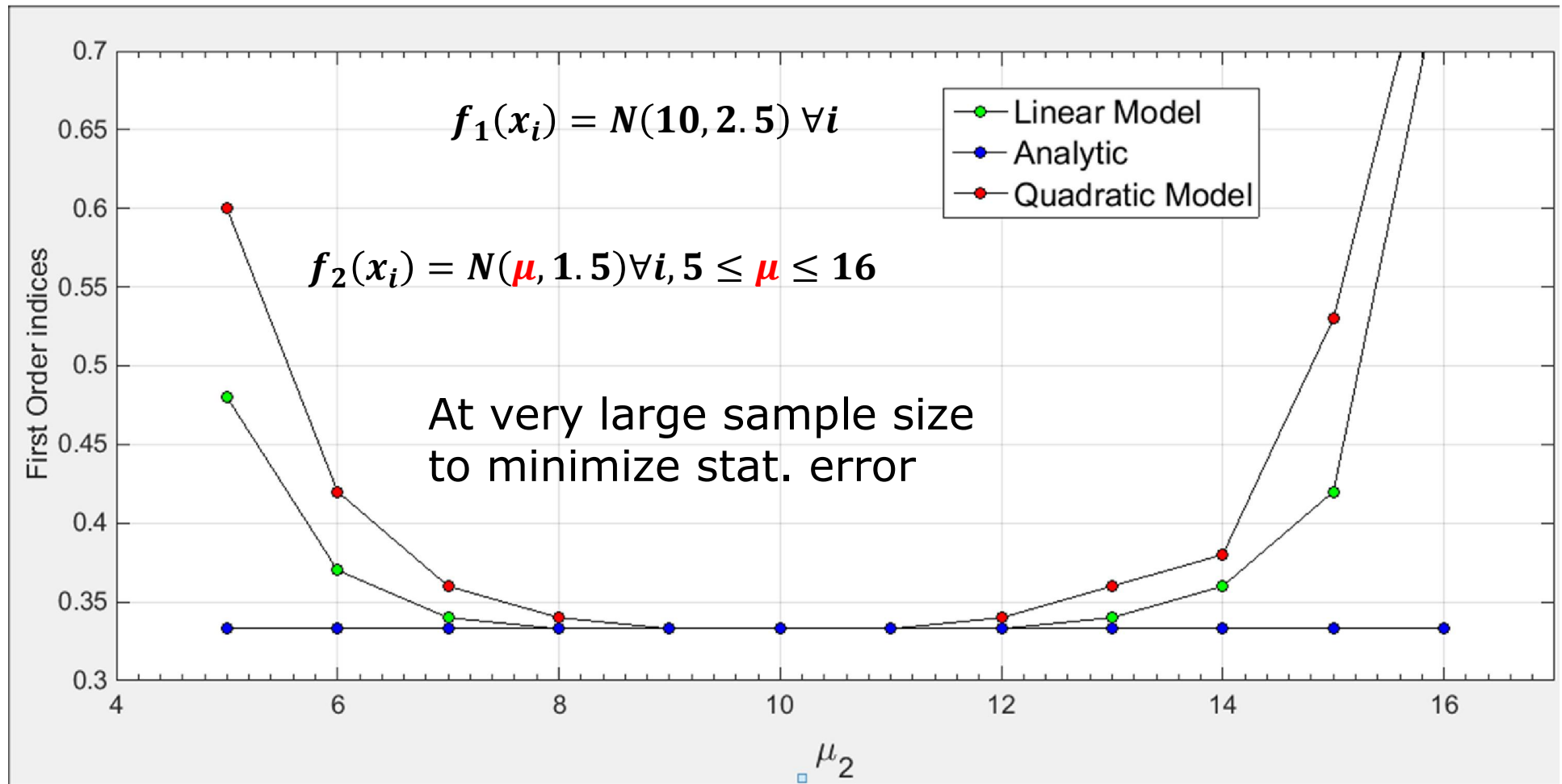
$$y = \sum_i^3 x_i$$

$$y = \sum_i^3 x_i^2$$

$$f_1(x_i) = N(10, 2.5) \forall i$$

$$f_2(x_i) = N(\mu, 1.5) \forall i, 5 \leq \mu \leq 16$$

Bias in Sensitivity Indices when F₂ shifts away from F₁





Conclusions

When the input pdf is changed, we usually re-run UA/SA using new model runs.

The approach presented here allows us to perform UA/SA using the same set of sample points (and model runs).

This is crucial when model runs are very CPU intensive.

The approach can manage both independent and dependent input.



Conclusions

An example shows that the approach performs better than the standard method of Beckman and McKay for uncertainty analysis.

The Kolmogorov-Smirnov test can be applied to see whether an update of the input pdf makes any changes in the output pdf.

We can repeat Sensitivity analysis on the new pdf without bias provided changes in the pdfs are limited.