EE6361 : Importance Sampling

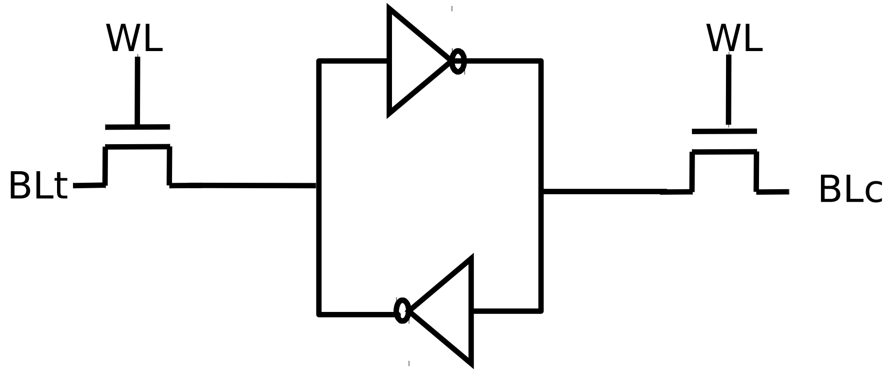
Electrical Engineering Department
IIT Madras

Janakiraman Viraraghavan

Agenda

- Yield analysis in memories
- Statistical Compact Model
- Traditional Monte-Carlo
- Limitations of Monte-Carlo
- Variance Reduction
- Importance Sampling
- Statistical Blockade

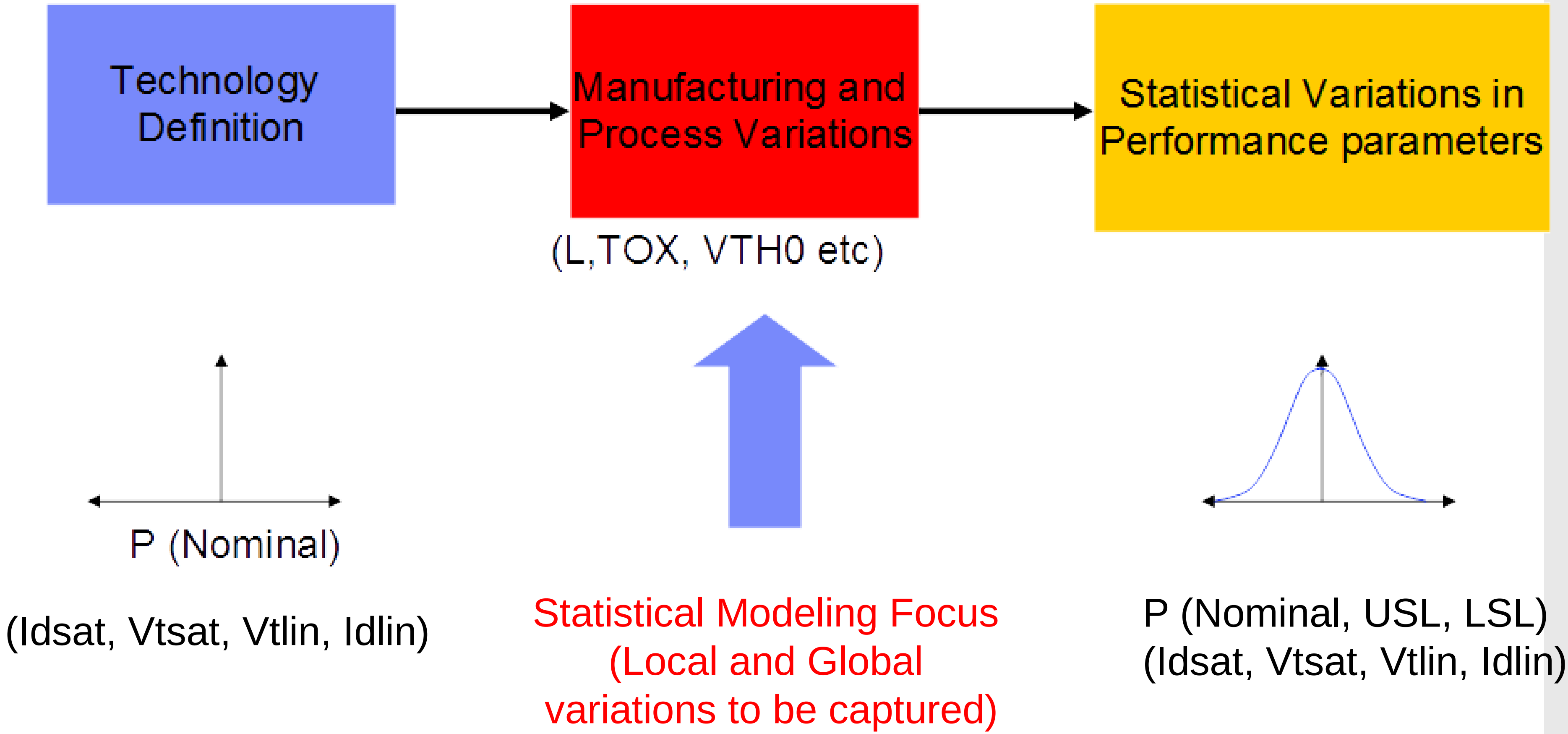
Yield Analysis in Memories



6T SRAM Bit Cell

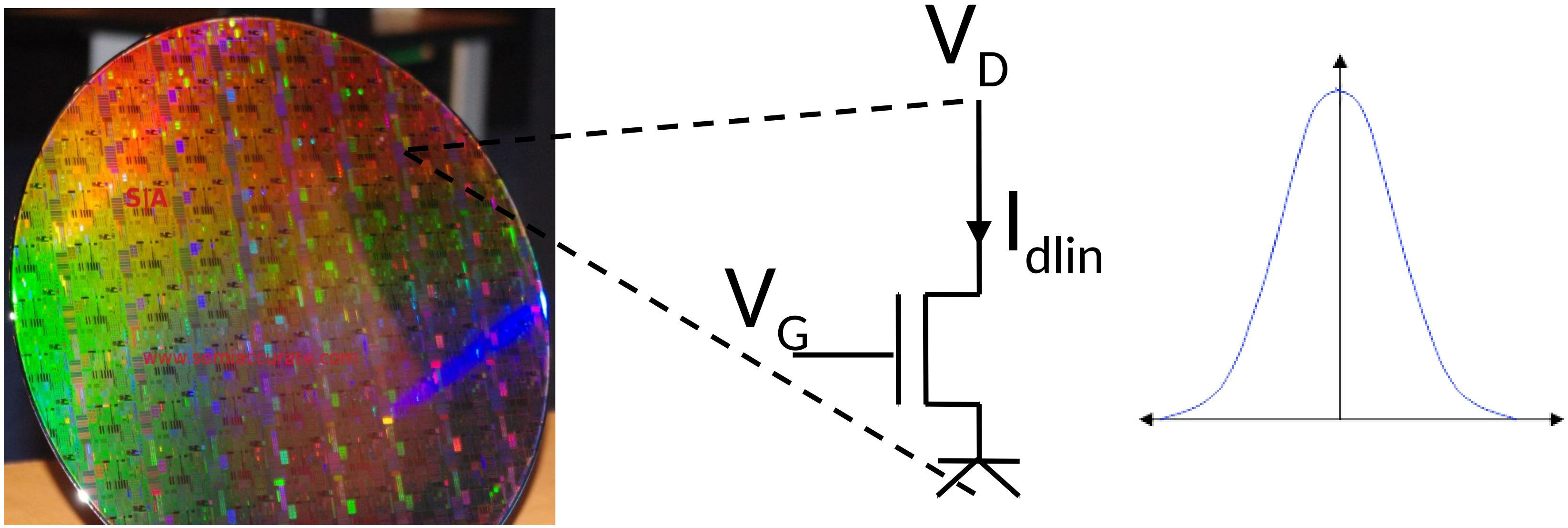
- Smallest possible transistors used
  - Prone to large variability
- Very high yield expected (99.9999%)
  - Foundry has to guarantee

Variability in Fab



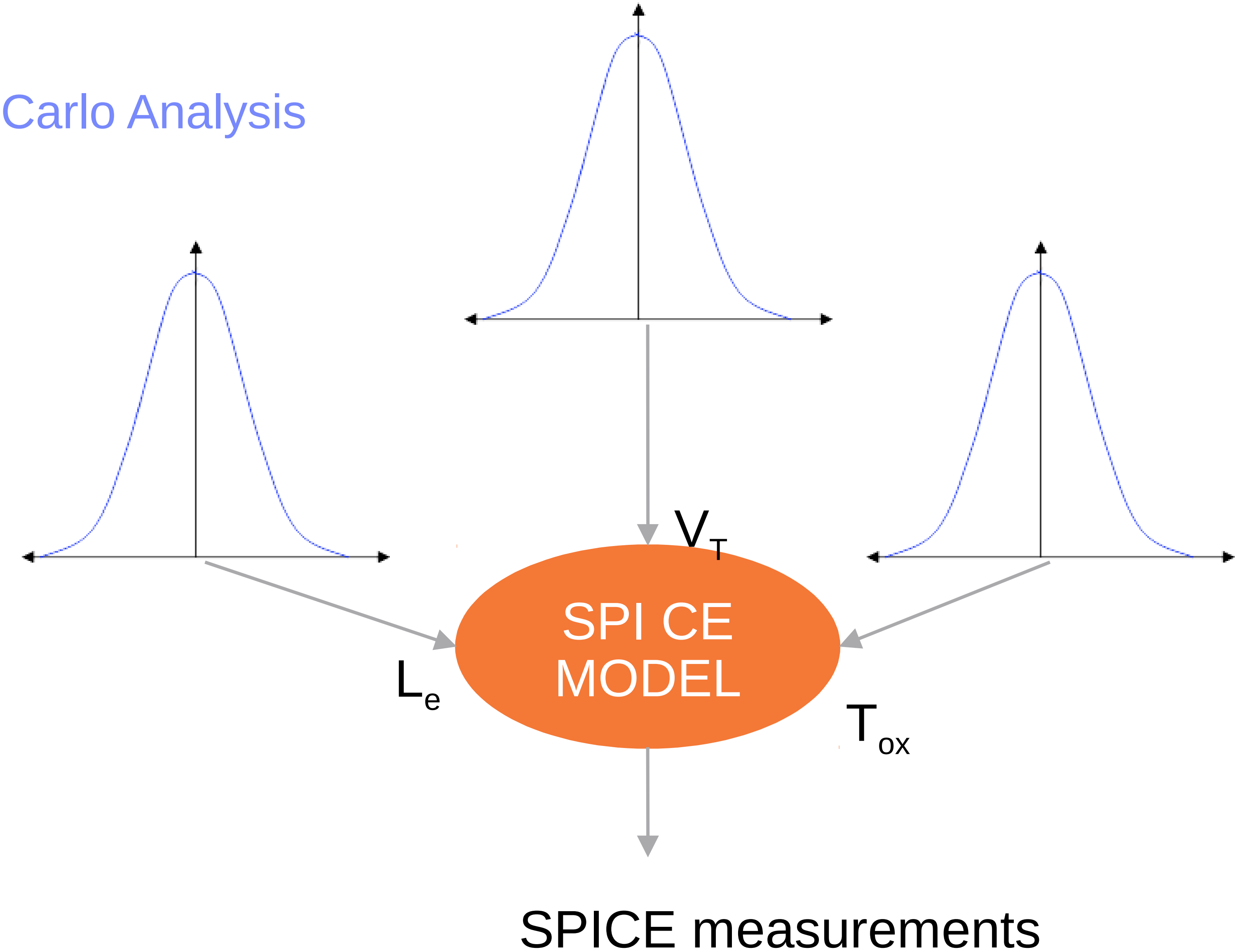
- Statistical compact model available to designers for use

Statistical Compact Model



- Maps measured variations of voltages and currents to process variations
  - Back propagation of Variance

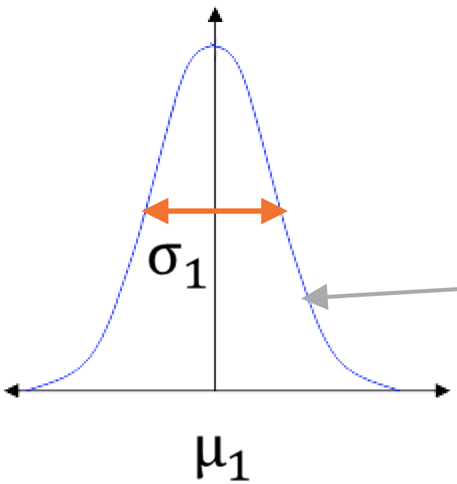
Monte Carlo Analysis



- Generate random samples from the distribution
- Simulate with each set of values

Monte Carlo Analysis

$$Y = h(X_1, X_2, \ldots X_K)$$



Trial #	$X_1, X_2 .. X_K$	Y
1	$x_1^1, x_2^1 ..... x_K^1$	Y^1
2	$x_1^2, x_2^2 ..... x_K^2$	Y^2
~		
N	$x_1^N, x_2^N ..... x_K^N$	Y^N
Mean		Sample Mean $\hat{Y}_1^N$
Variance		

Monte Carlo Analysis

$$\hat{Y}_N = \frac{1}{N}\sum_{i=1}^{N} Y_i$$

Monte Carlo Run 1

Monte Carlo Run 1

Monte Carlo Run 1

Monte Carlo Run M		
Trial #		
1		
2		
~		
N		
Mean		$\hat{Y}_M^N$
Variance		

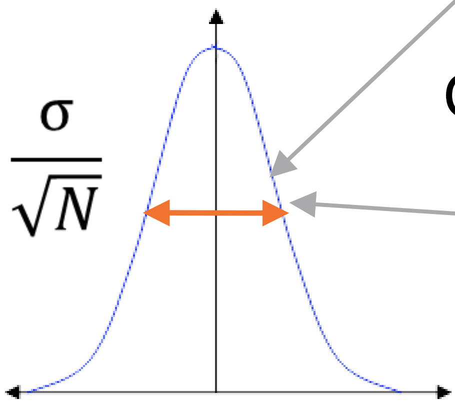
Run #	Mean
1	$\hat{Y}_1^N$
2	$\hat{Y}_1^N$
~	
M	$\hat{Y}_M^N$

Monte Carlo Analysis

$$\hat{Y}_N = \frac{1}{N}\sum_{i=1}^{N} Y_i$$

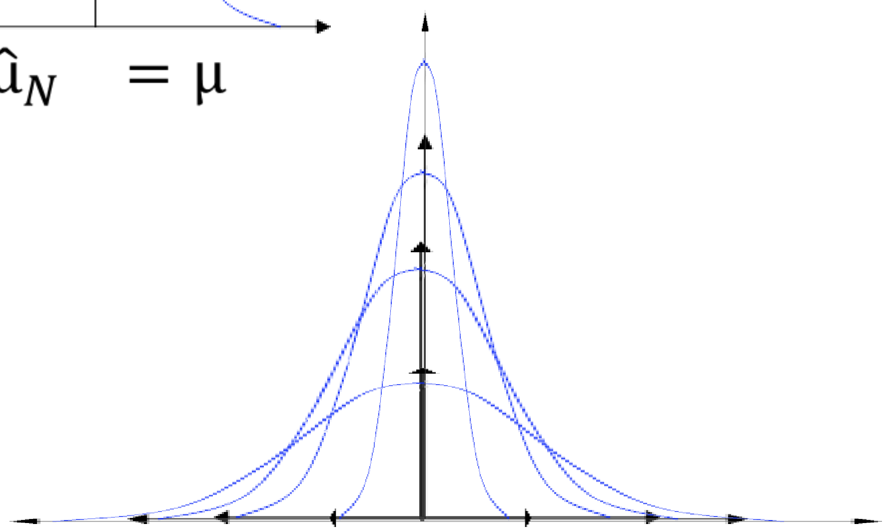
CLT - Gaussian

$$\hat{\sigma}_N = \frac{\sigma}{\sqrt{N}}$$



$$\hat{\mu}_N = \mu$$

Run #	Mean
1	$\hat{Y}_1^{1}$
2	$\hat{Y}_1^{2}$
~	
M	$\hat{Y}_1^{M}$

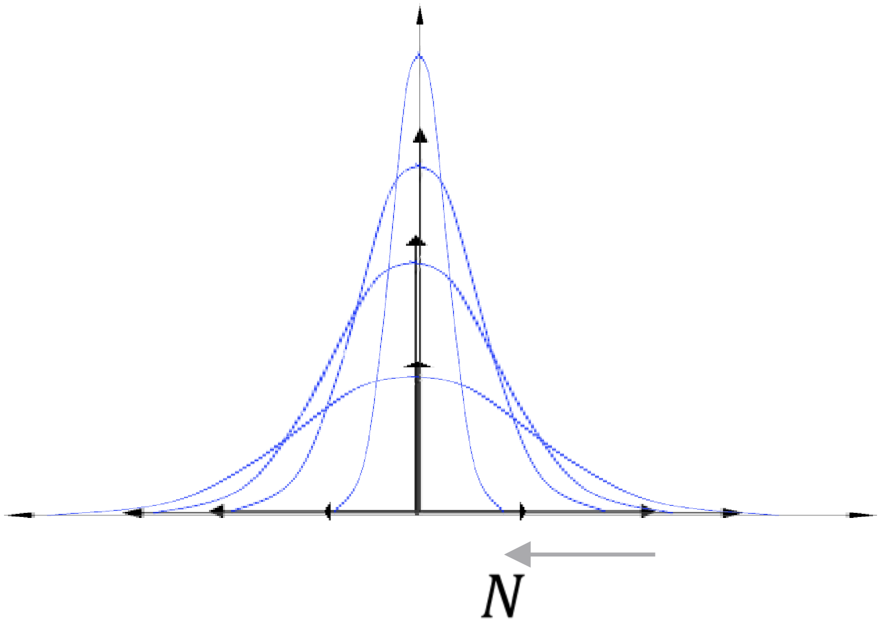


Monte Carlo Analysis – Variance Reduction

$$\hat{Y}_N = \frac{1}{N}\sum_{i=1}^{N} Y_i$$

$$\hat{\sigma}_N = \frac{\sigma}{\sqrt{N}}$$

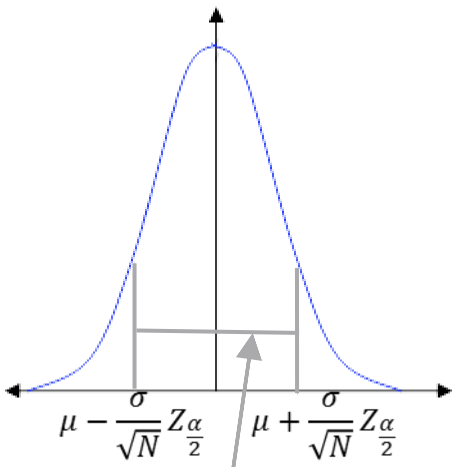
$$\hat{\mu}_N = \mu$$



- As N increases variance decreases
- By increasing N, you can get arbitrarily close

Monte Carlo Analysis – Estimating N

$$P\left(\frac{\left|\hat{Y}_N - \mu\right|}{\frac{\sigma}{\sqrt{N}}} > Z_{\frac{\alpha}{2}}\right) = 1 - \alpha$$



Standard Normal RV

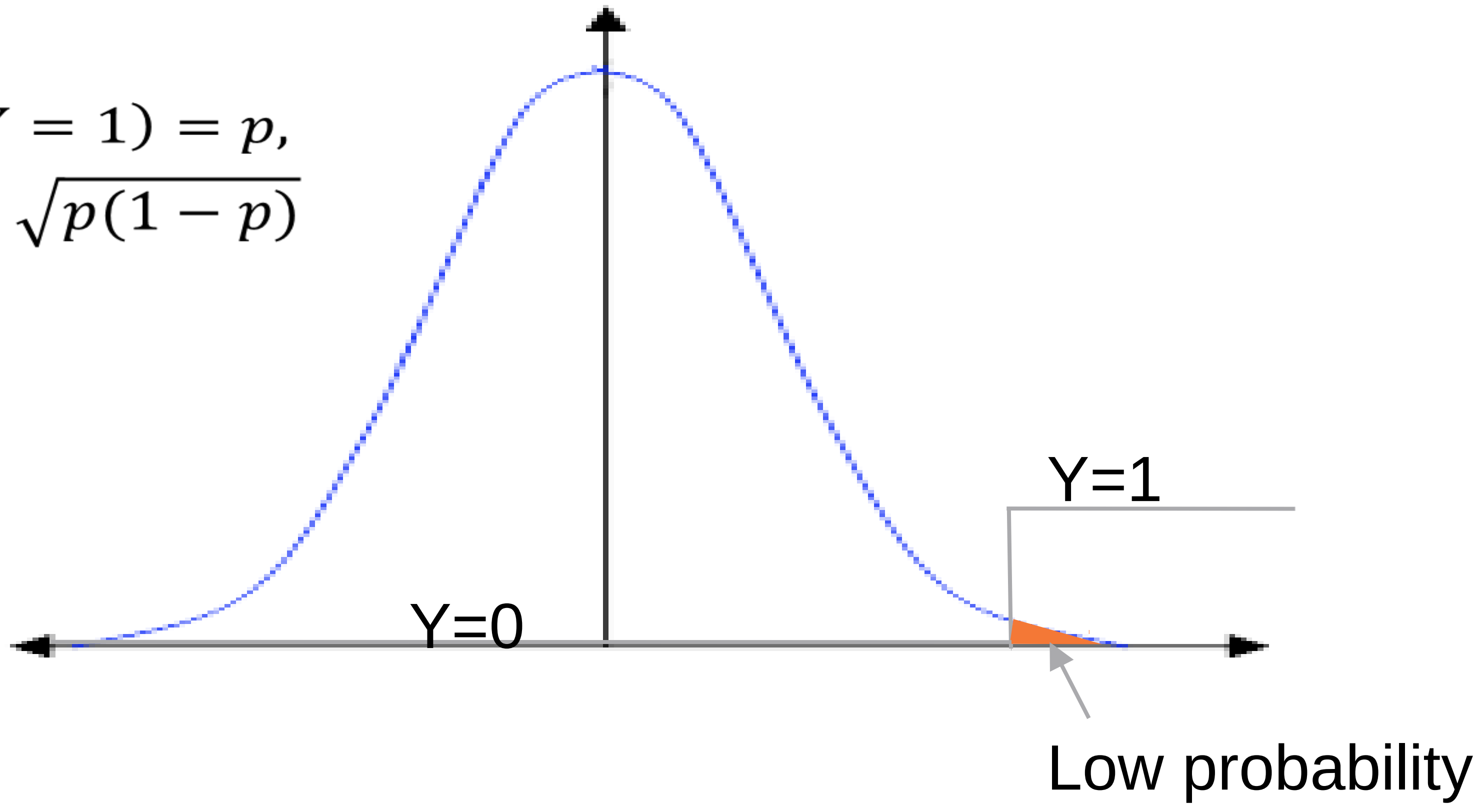
Confidence interval $= \epsilon$

$$2\frac{\sigma}{\sqrt{N}} Z_{\frac{\alpha}{2}} = \epsilon$$

$$N \geq \left[2\left(\frac{\sigma}{\epsilon}\right) Z_{\frac{\alpha}{2}}\right]^2$$

Rare Event Estimation

$$P(Y = 1) = p,$$
$$\sigma = \sqrt{p(1-p)}$$



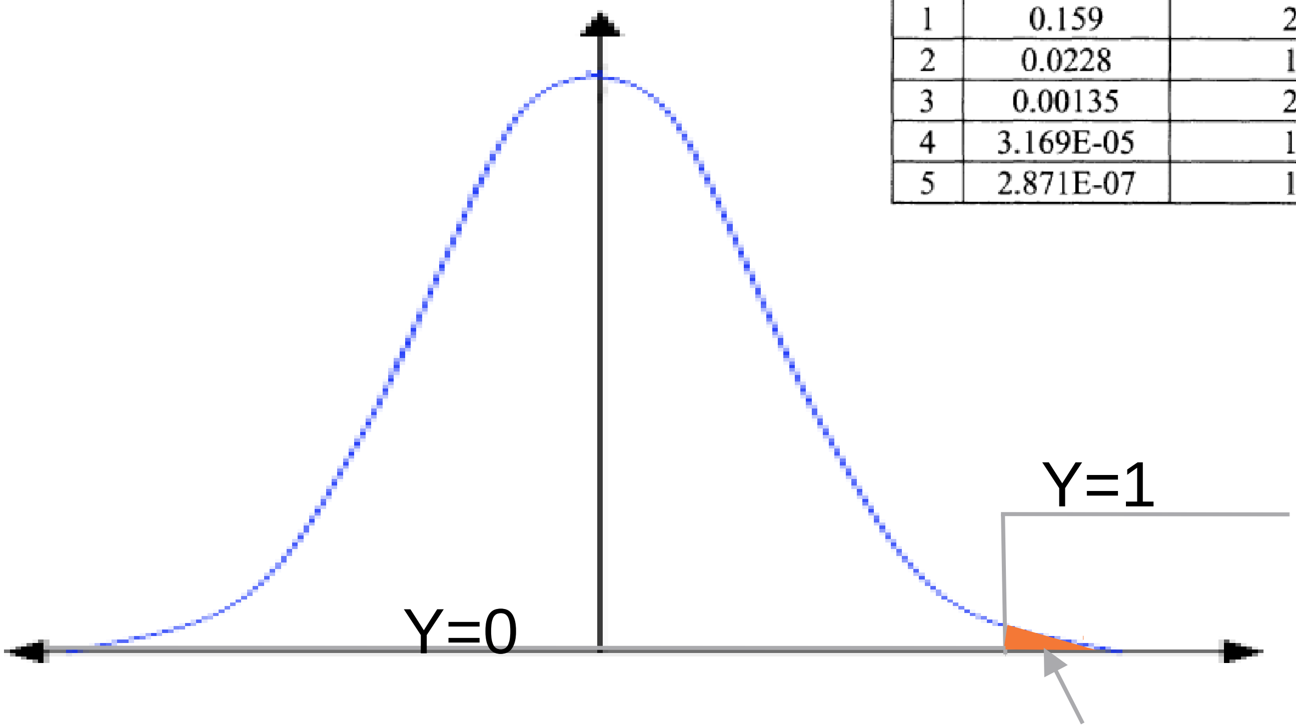
$$\varepsilon = 2\alpha p$$
$$Z_{\frac{\alpha}{2}} = 2 \; - 95\% \; confidence$$

$$N \geq \left[\frac{4}{\alpha^2}\left(\frac{1-p}{p}\right)\right]$$

Monte Carlo Analysis – Estimating N

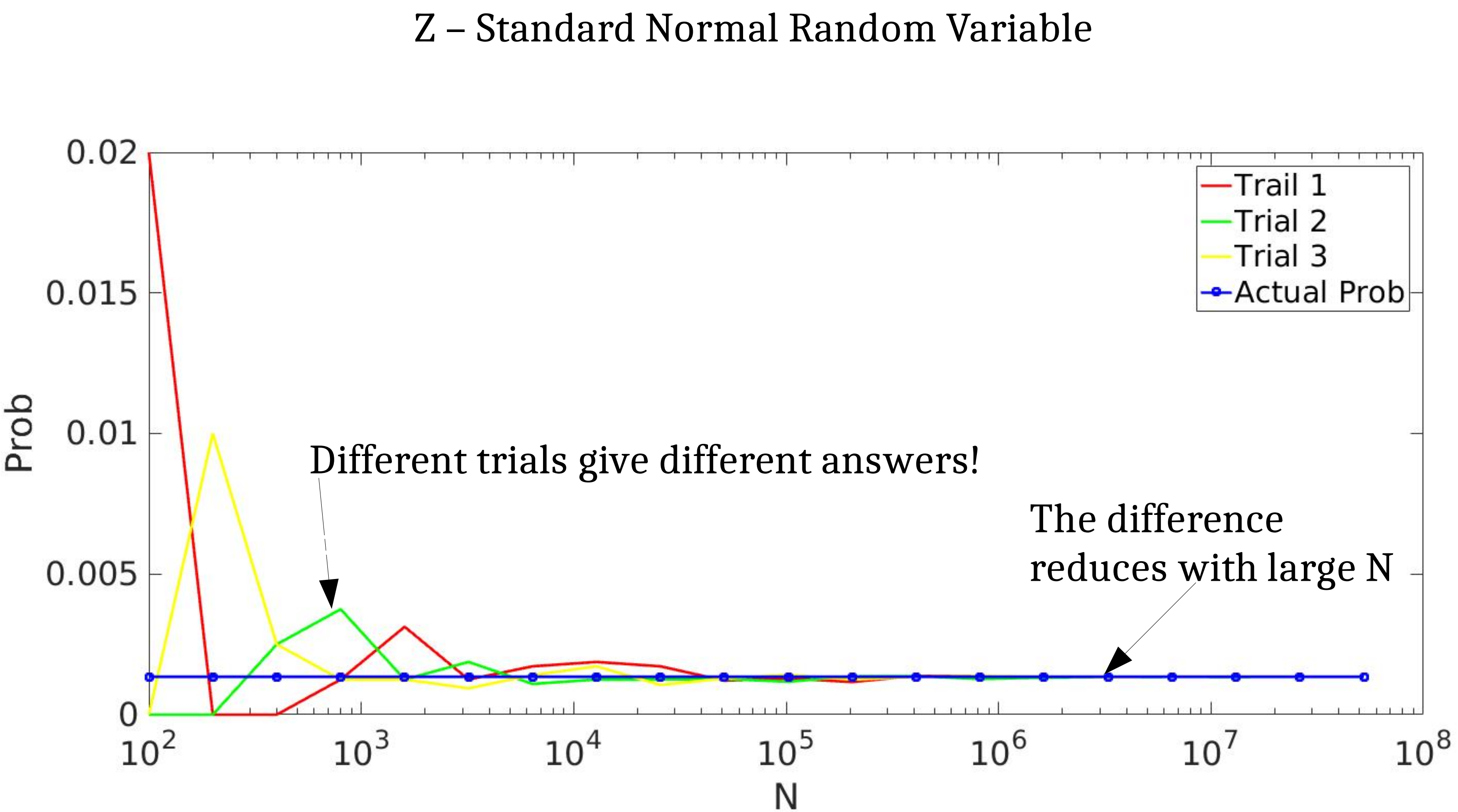
Table 1. Number of Monte Carlo simulations needed to estimate the probability P_f=Prob(x>z_0) with a 95% confidence interval = [P_f-0.1P_f, P_f+0.1P_f]. The corresponding simulation runtime in days, if a spice-like simulator is used.

z_0	Probability value	Number of Monte Carlo Simulations	Runtime in Number of Days
0	0.500	4.0e2	0.18
1	0.159	2.1e3	0.972
2	0.0228	1.7e4	7.87
3	0.00135	2.9e5	Too long!
4	3.169E-05	1.3e7	Too long!
5	2.871E-07	1.4e9	Too long !



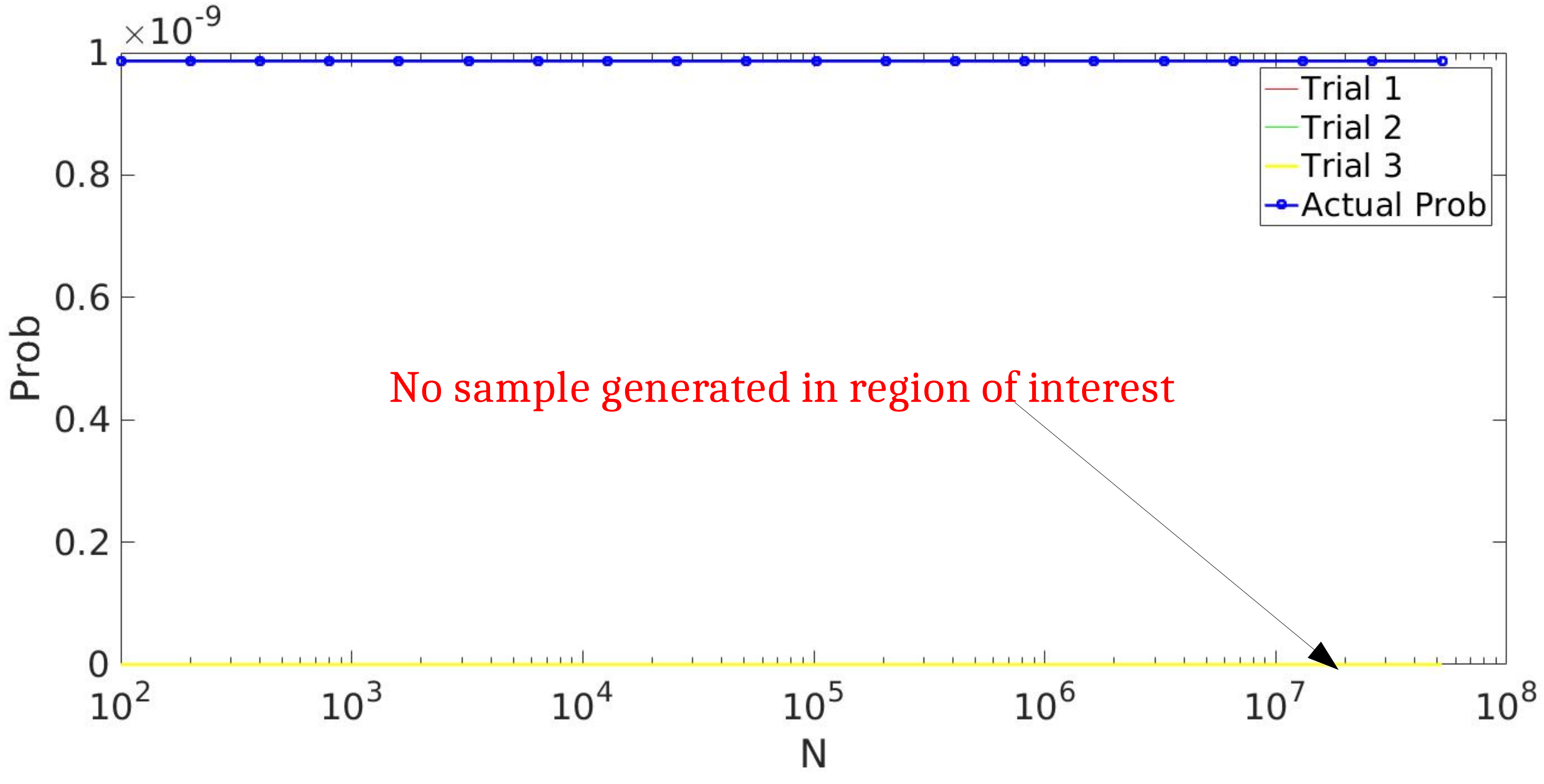
Rare event – "Unlikely" to be sampled by Monte-Carlo

Monte Carlo Estimation P(Z > 1.5)



Monte Carlo Estimation P(Z > 6)

Z – Standard Normal Random Variable

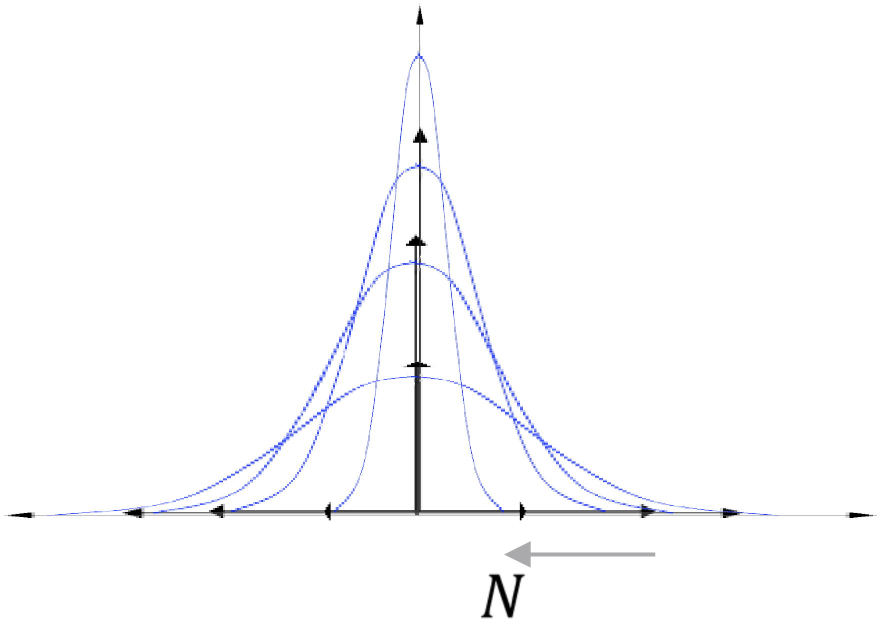


Monte Carlo Analysis – Variance Reduction

$$\hat{Y}_N = \frac{1}{N}\sum_{i=1}^{N} Y_i$$

$$\hat{\sigma}_N = \frac{\sigma}{\sqrt{N}}$$

$$\hat{\mu}_N = \mu$$



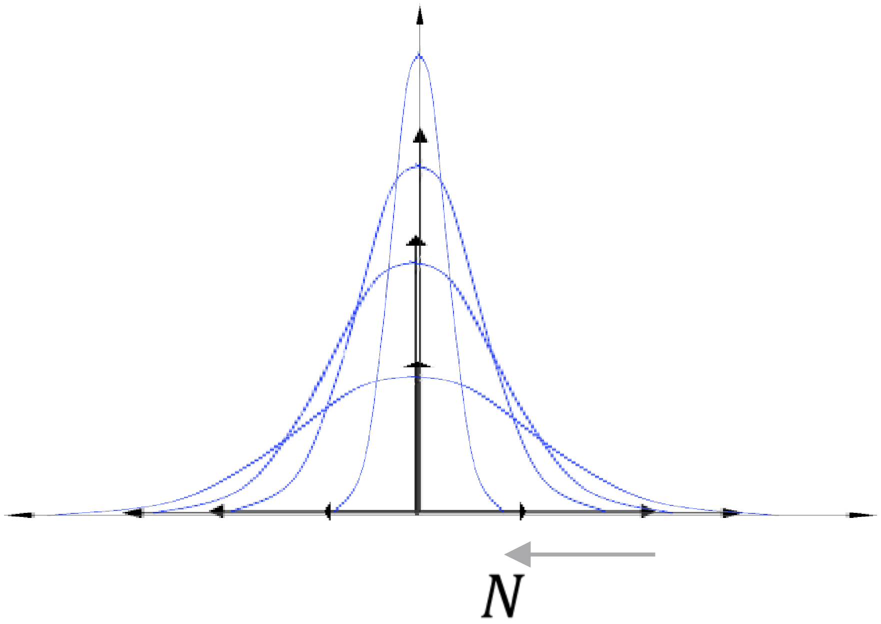
- As N increases - variance decreases
- Can you reduce the variance in any other way?

Monte Carlo Analysis – Variance Reduction

$$\hat{Y}_N = \frac{1}{N}\sum_{i=1}^{N} Y_i$$

$$\hat{\sigma}_N = \frac{\sigma}{\sqrt{N}}$$

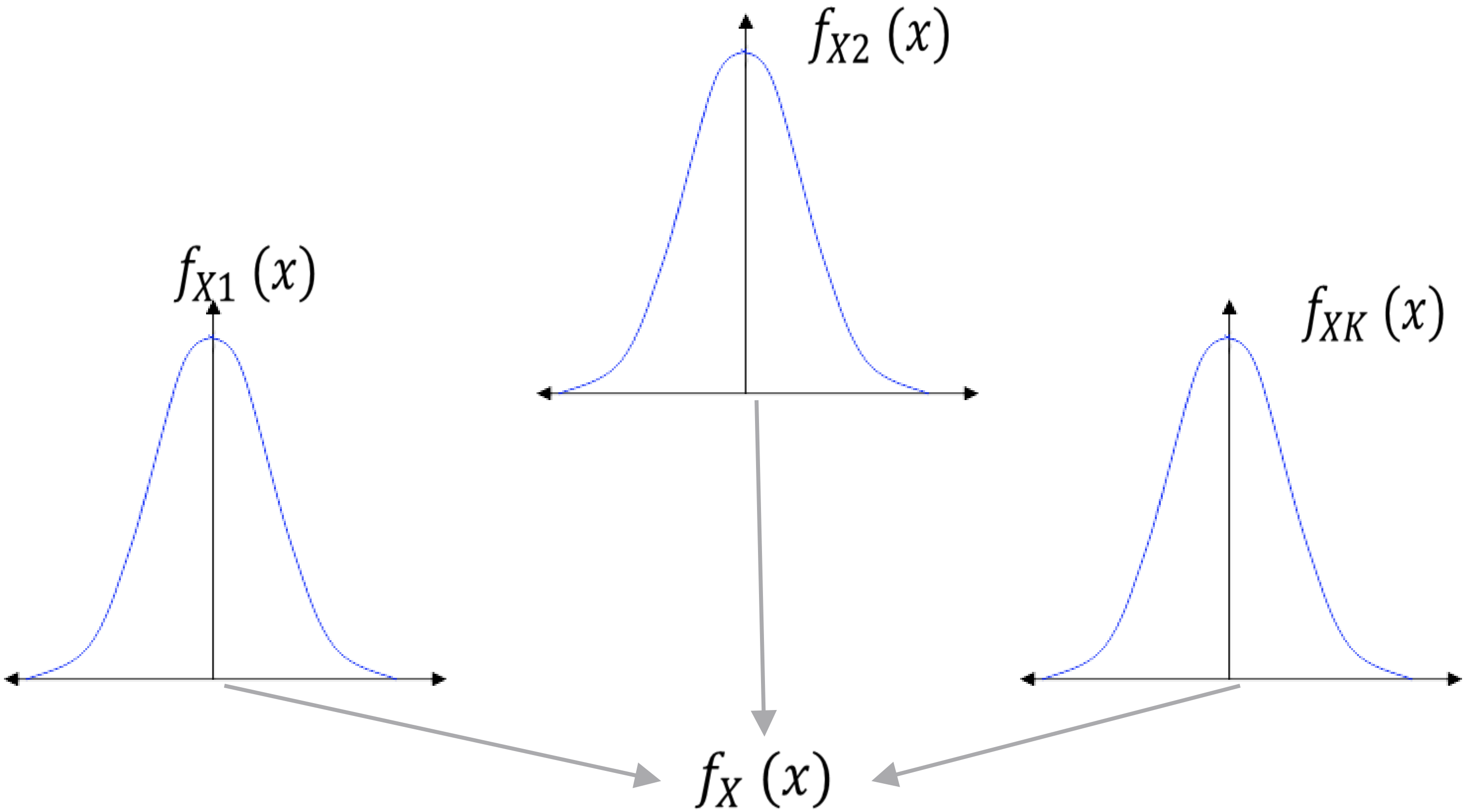
$$\hat{\mu}_N = \mu$$



- As N increases - variance decreases
- Can you reduce the variance in any other way?

Monte Carlo Analysis – Variance Reduction

$$Y = h(X_1, X_2, \dots X_K)$$



Monte Carlo Analysis – Variance Reduction

$$Y = h(X_1, X_2, \ldots X_K)$$

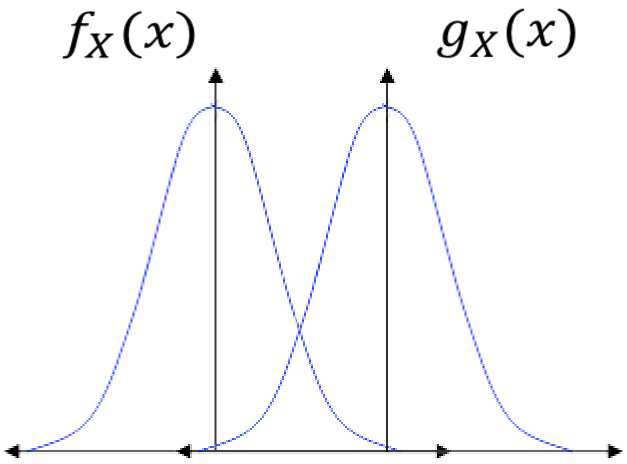
$$\mu = E_f[Y] = \int h(x) f_X(x) dx$$

$$\sigma^2 = Var_f[Y] = E_f[Y^2] - (E_f[Y])^2$$

$$\int (h(x))^2 f_X(x) dx$$

Importance Sampling

$$Y = h(X_1, X_2, \dots X_K)$$



$$Z = h(X)\frac{f_X(X)}{g_X(X)}$$

Define a new function Z where $g_X(x)$ is a new distribution
Z should have the same mean as Y by lower variance

Importance Sampling - Mean

Y - X is drawn from f()	Z – X is drawn from g()
$E_f[Y] = \int h(x) f_X(x) dx$	$E_g[Z] = \int h(x) \frac{f_X(x)}{\cancel{g_X(x)}} \cancel{g_X(x)} dx$

Importance Sampling - Variance

Y - X is drawn from f()	Z – X is drawn from g()
$Y = h(X)$	$Z = h(X)\frac{f_X(X)}{g_X(X)}$
$Var_f[Y] = E_f[Y^2] - (E_f[Y])^2$	$Var_g[Z] = E_g[Z^2] - (E_g[Z])^2$
$E_f[Y] = \int h(x) f_X(x) dx$	$E_g[Z] = \int h(x)\frac{f_X(x)}{g_X(x)} g_X(x) dx$
$E_f[Y^2] = \int (h(x))^2 f_X(x) dx$	$E_g[Z^2] = \int (h(x))^2 \left(\frac{f_X(x)}{g_X(x)}\right)^2 g_X(x) dx$

Importance Sampling - Variance

$$E_g[Z^2] < E_f[Y^2]$$

Desired!

$$E_f[Y^2] = \int (h(x))^2 f_X(x)dx$$

$$E_g[Z^2] = \int (h(x))^2 \left(\frac{f_X(x)}{g_X(x)}\right)^2 g_X(x)dx$$

$$E_g[Z^2] = \int (h(x))^2 \left(\frac{f_X(x)}{g_X(x)}\right) f_X(x)dx$$

Importance Sampling - Variance

$$E_g[Z^2] < E_f[Y^2]$$

Desired!

$$E_f[Y^2] = \int (h(x))^2 f_X(x)dx$$

$$E_g[Z^2] = \int (h(x))^2 \left(\frac{f_X(x)}{g_X(x)}\right) f_X(x)dx$$

Choose g(x) such that this ratio less than 1 for all values of x!

Importance Sampling – Choice of g(x)

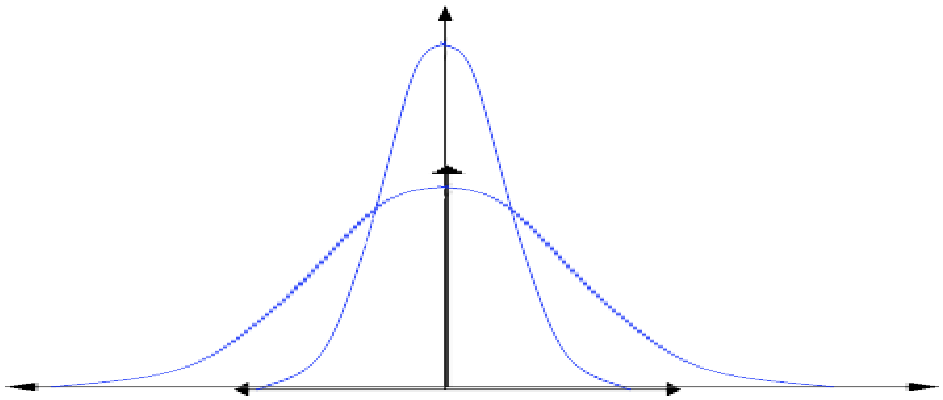
$$E_g\left(\frac{f_X(x)}{g_X(x)}\right) = \int \left(\frac{f_X(x)}{g_X(x)}\right) g_X(x)dx \ = \ 1$$

How can a function be less than 1 for all values of x
But have mean of 1?

Importance Sampling – Variance – Choice of g(x)

$$E_g\left(\frac{f_X(x)}{g_X(x)}\right) = \int \left(\frac{f_X(x)}{g_X(x)}\right) g_X(x)dx \ = \ 1$$

How can a function be less than 1 for all values of x
But have mean of 1?



g(x) is also a PDF – Has to integrate to unity!

Importance Sampling – Choice of g(x)

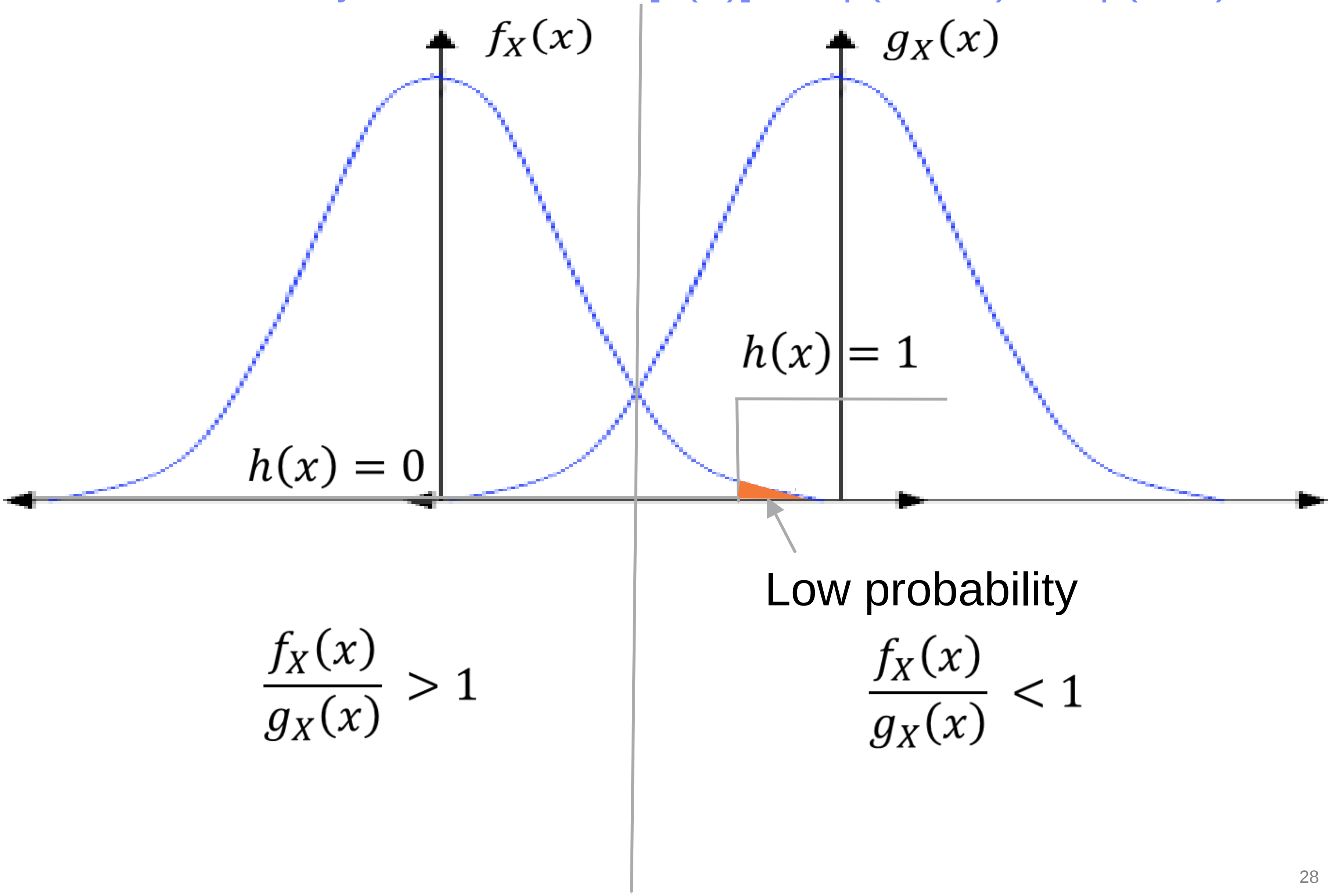
$$E_g[Z^2] < E_f[Y^2]$$

$$E_f[Y^2] = \int (h(x))^2 f_X(x)dx$$

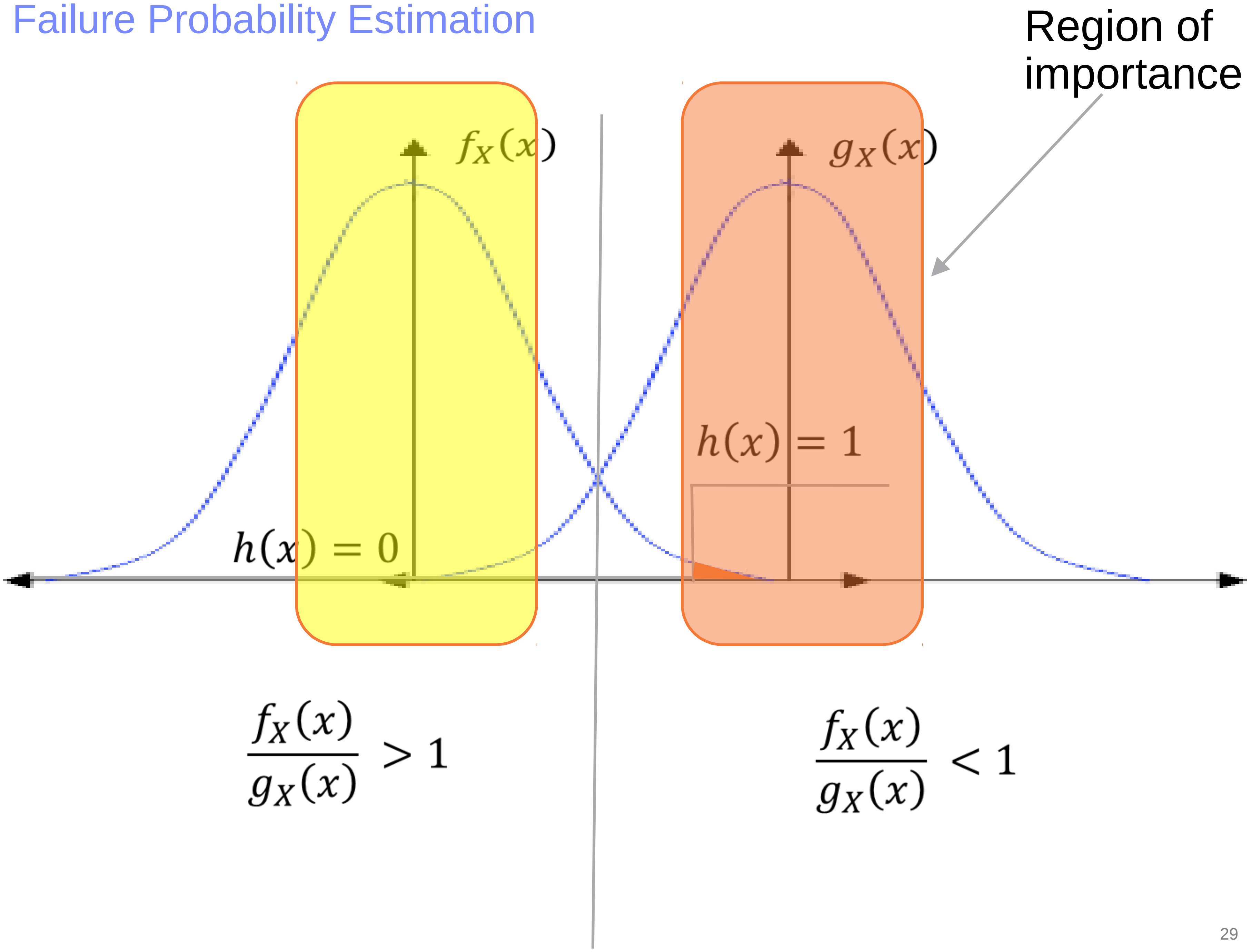
$$E_g[Z^2] = \int (h(x))^2 \left(\frac{f_X(x)}{g_X(x)}\right) f_X(x)dx$$

Ratio will be greater than 1 for some values of x
If h(x) is strictly zero in those regions?

Failure Probability Estimation – E[h(X)] = 1*p(X>= a) + 0*p(X<a)



Failure Probability Estimation

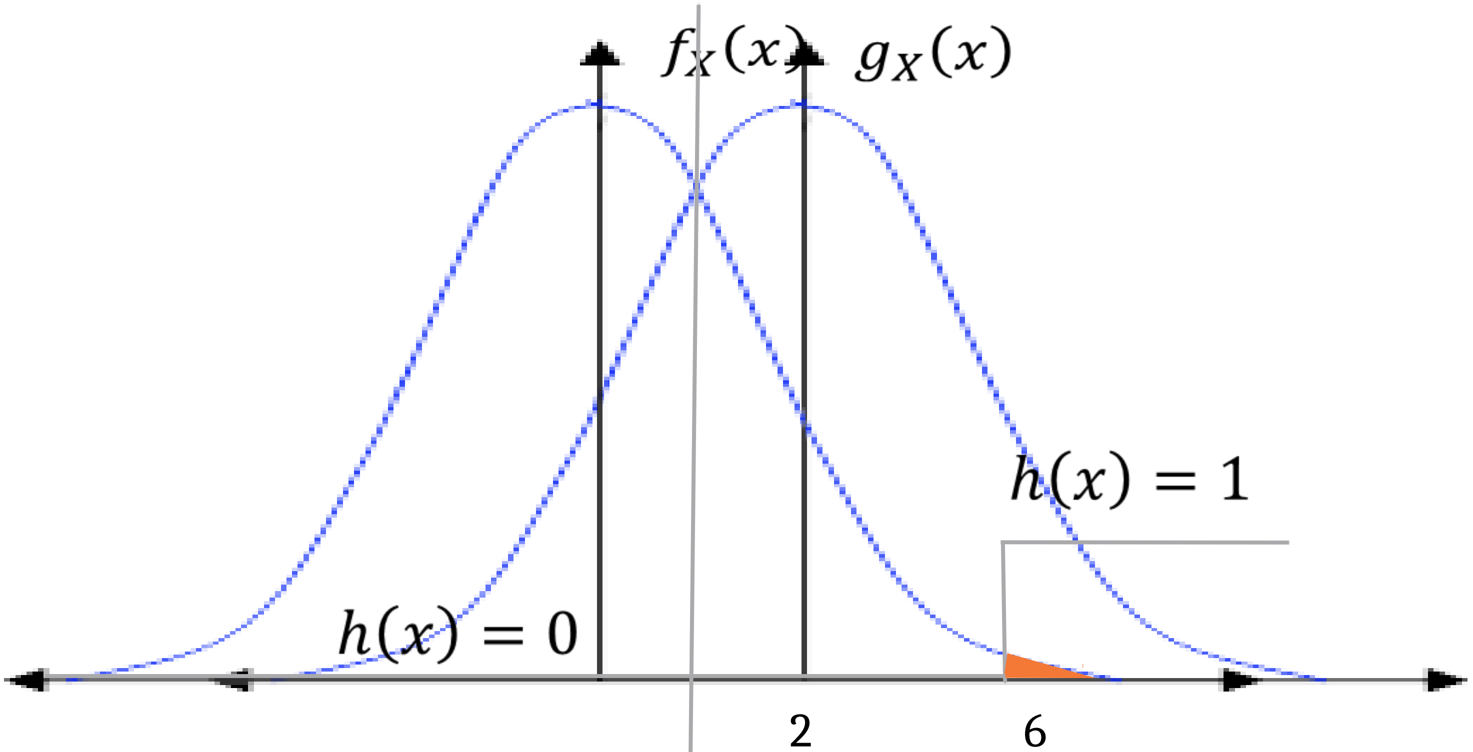


$$\frac{f_X(x)}{g_X(x)} > 1$$

$$\frac{f_X(x)}{g_X(x)} < 1$$

Example: P(Z>6)

Z – Standard Normal Random Variable

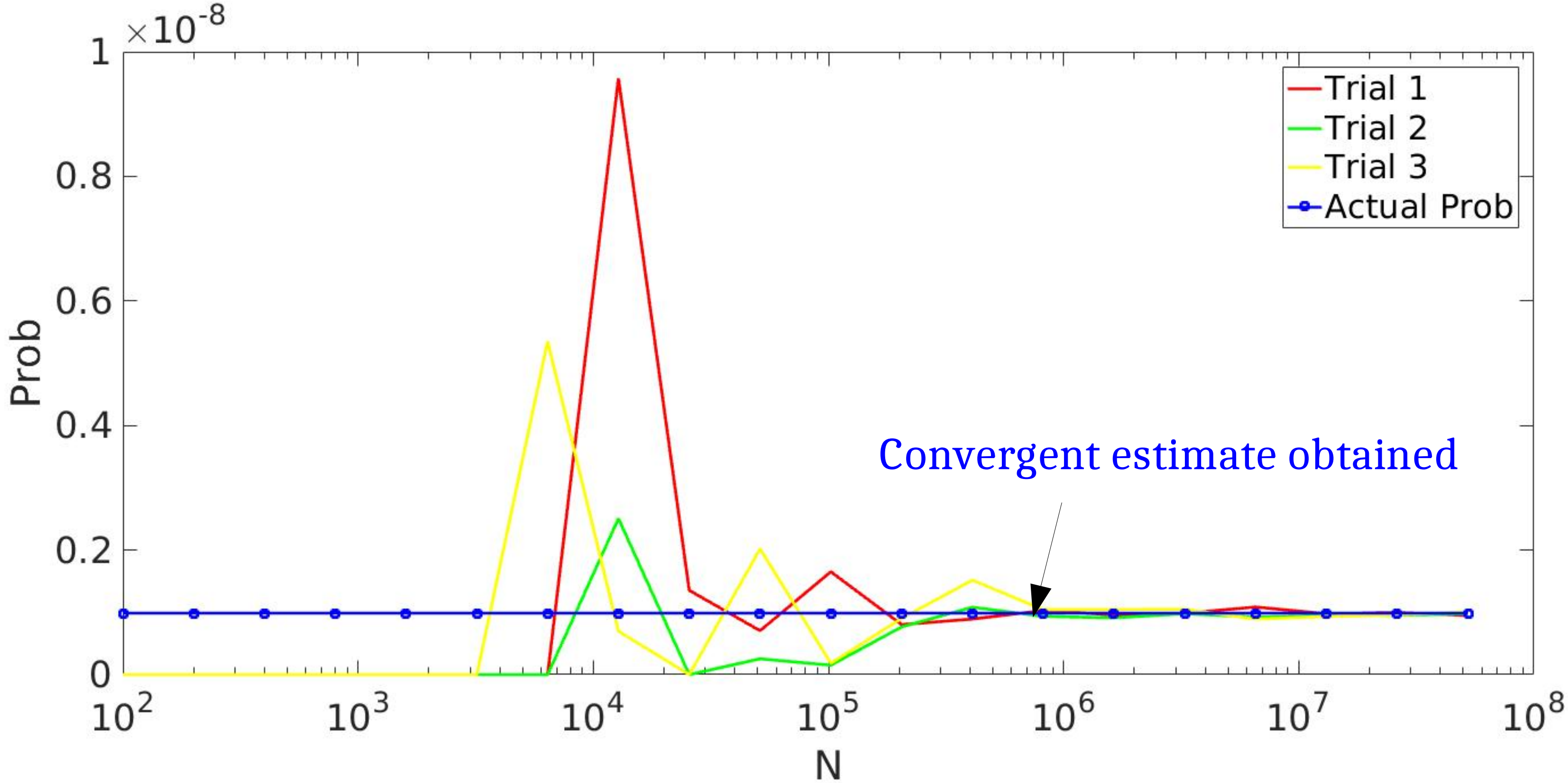


$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}$$

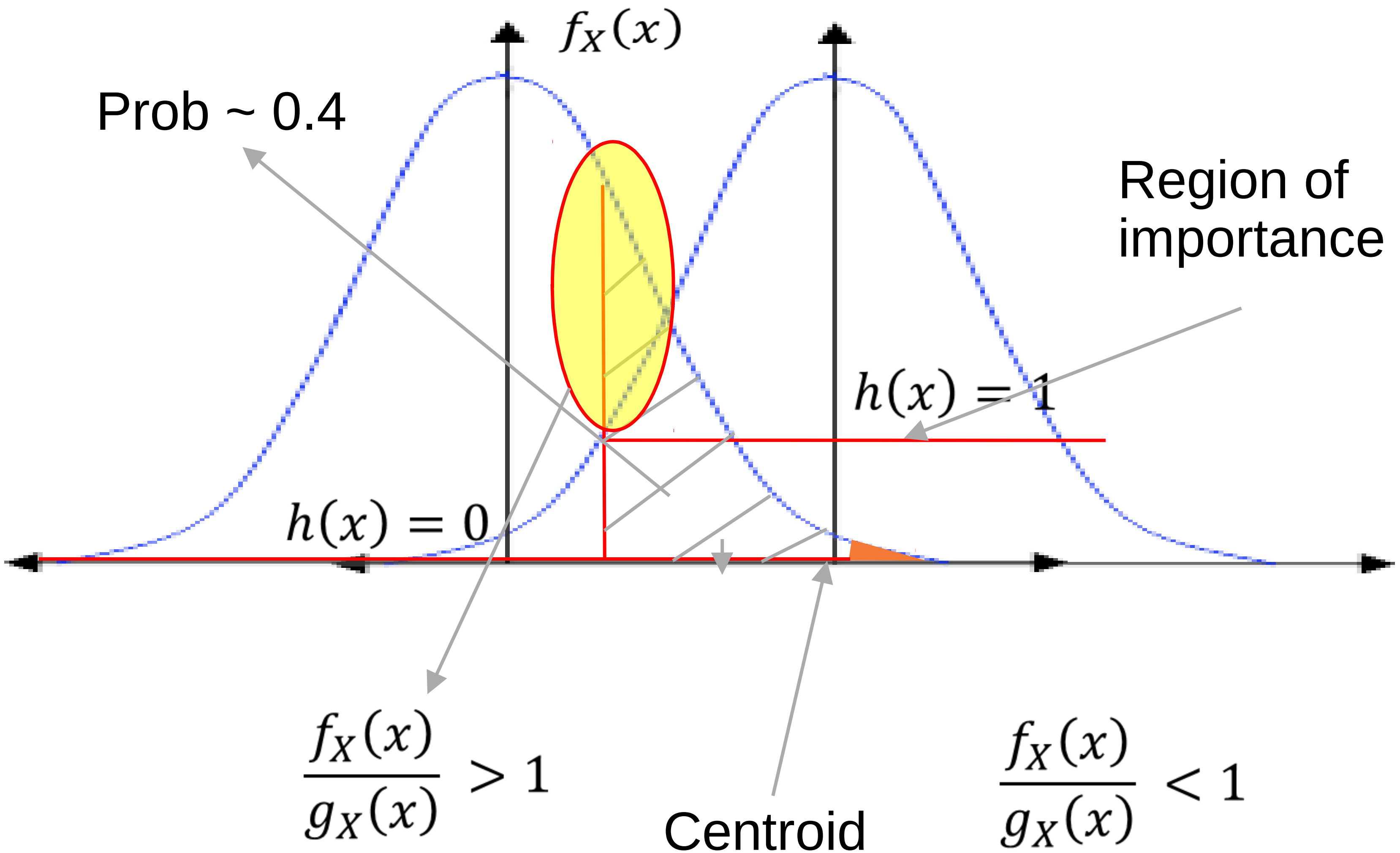
$$g_X(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-(x-2)^2}{2}}$$

Estimation with Importance Sampling P(Z > 6)

Z – Standard Normal Random Variable



Failure Probability Estimation - Where does it fail?



Beware when estimating large probabilities!

References

- Kanj, R.; Joshi, R.; Nassif, S., "Mixture importance sampling and its application to the analysis of SRAM designs in the presence of rare failure events," in *Design Automation Conference, 2006 43rd ACM/IEEE* , vol., no., pp.69-72, 0-0 0 doi: 10.1109/DAC.2006.229167
- “Probability Models” second edition by Sheldon Ross
- “Introduction to Probability and Statistics” by Sheldon Ross