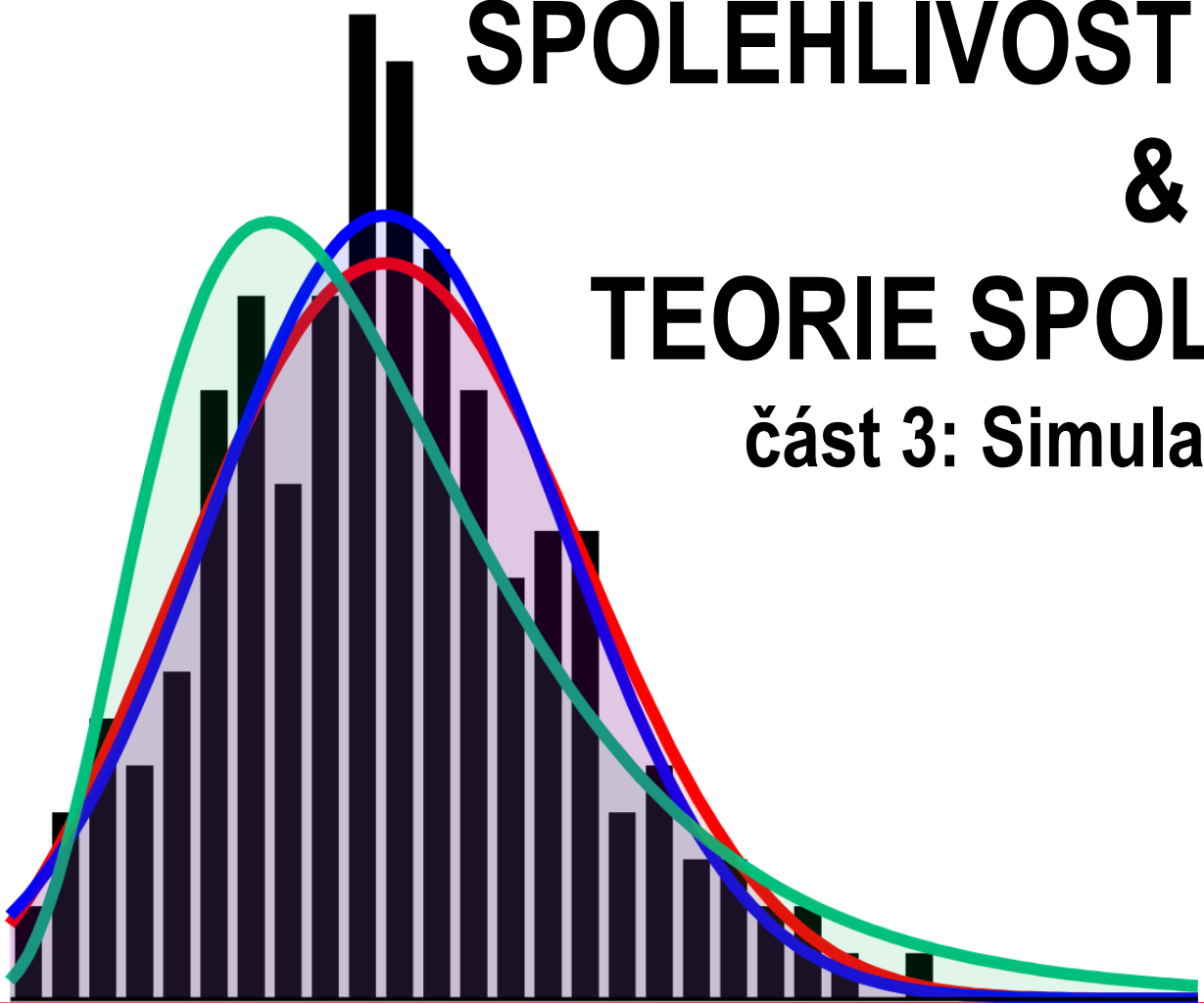




# SPOLEHLIVOST KONSTRUKCÍ & TEORIE SPOLEHLIVOSTI

## část 3: Simulační metody



Drahomír Novák  
Jan Eliáš



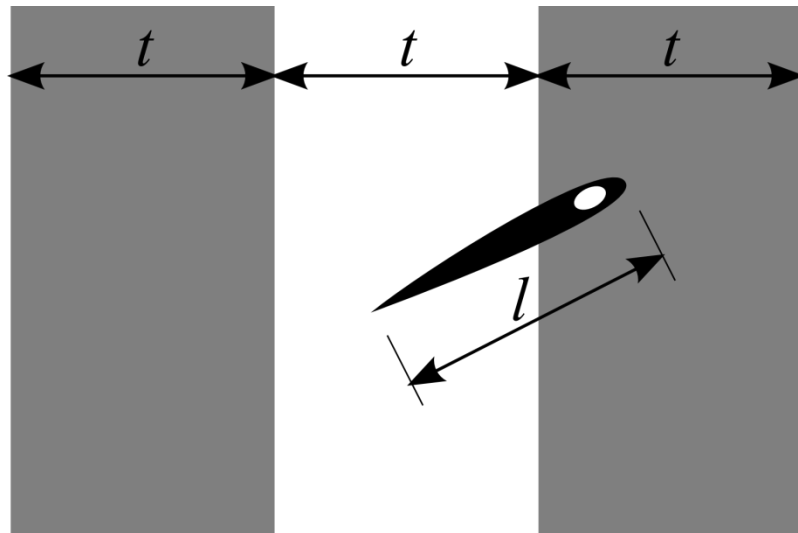
# část 3 Simulační metody

# Buffonova jehla (18. století)

- přímky o vzdálenosti  $t$
- pravděpodobnost** že jehla délky  $l$  překříží přímku je (pokud  $l \leq t$ )



$$p = \frac{2l}{t\pi}$$



Obrázek 2. Jehla křížící přímku

- házel jehlu  $N_{sim}$  krát, pravděpodobnost  $p$  je počet hodů s překřížením dělený  $N_{sim}$
- umožňuje to odhad  $\pi$

Obrázek 1. Georges-Louis Leclerc, Comte de Buffon



# Metoda Monte Carlo

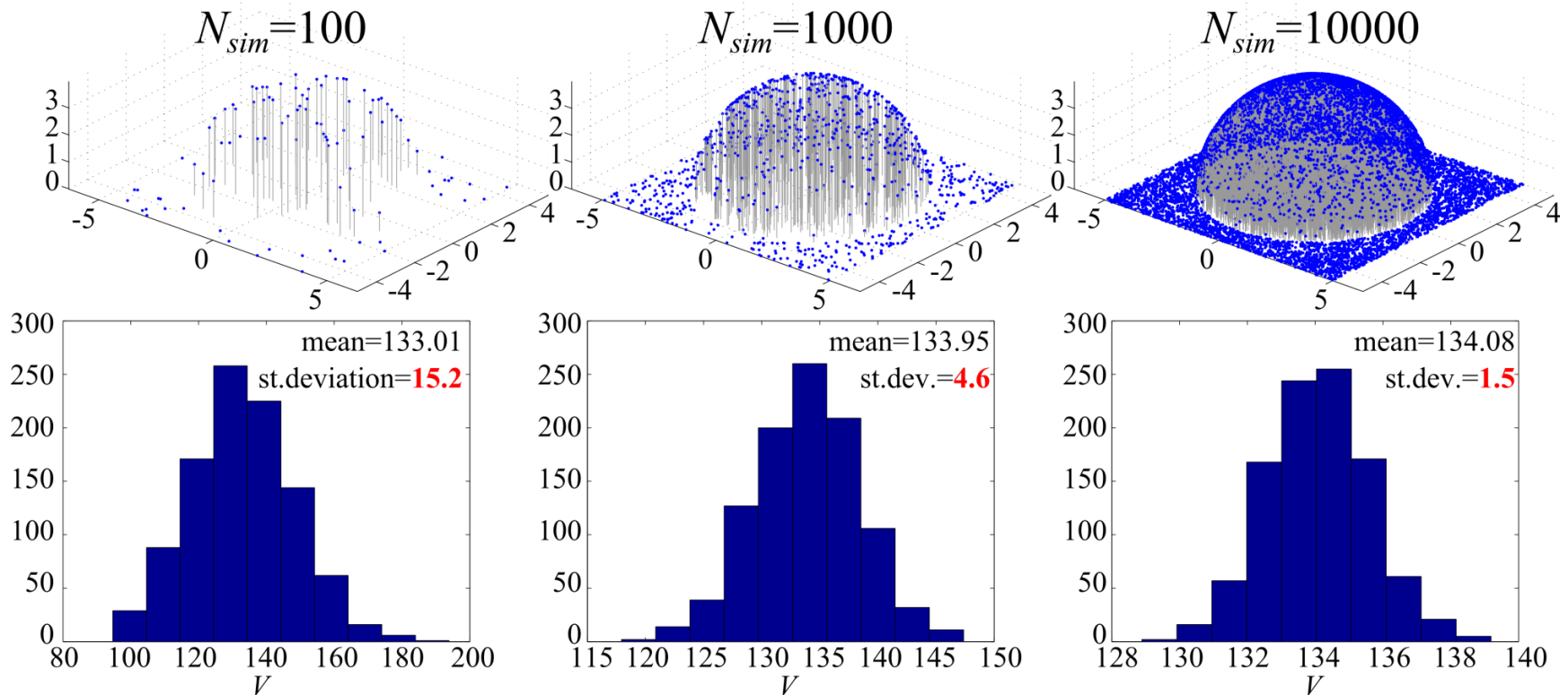
- metoda Monte Carlo poskytuje **přibližné řešení** rozmanitých matematických problému prováděním statistických experimentů na počítači
- chyba klesá lineárně s  $\sqrt{1/N_{sim}}$ , kde  $N_{sim}$  je počet simulací (virtuálních experimentů)
- například se používá pro **numerické odhady integrálů**
- generuje  $N_{sim}$  vstupních hodnot  $\mathbf{x}_i$  v objemu  $V_{sim}$
- numerická integrace se pak provede následovně

$$\int_{V_{sim}} f(\mathbf{x})d\mathbf{x} = \frac{V_{sim}}{N_{sim}} \sum_{i=1}^{N_{sim}} f(\mathbf{x}_i)$$



# Příklad integrace metodou MC

- objem polokoule – funkce  $z = \sqrt{4^2 + x^2 - y^2}$
- body generovány na oblasti  $x \in \langle -5,5 \rangle$   $y \in \langle -5,5 \rangle$ ,  $V_{sim} = 100$
- teoretické řešení  $V = 4/6\pi r^3 = 134.04$



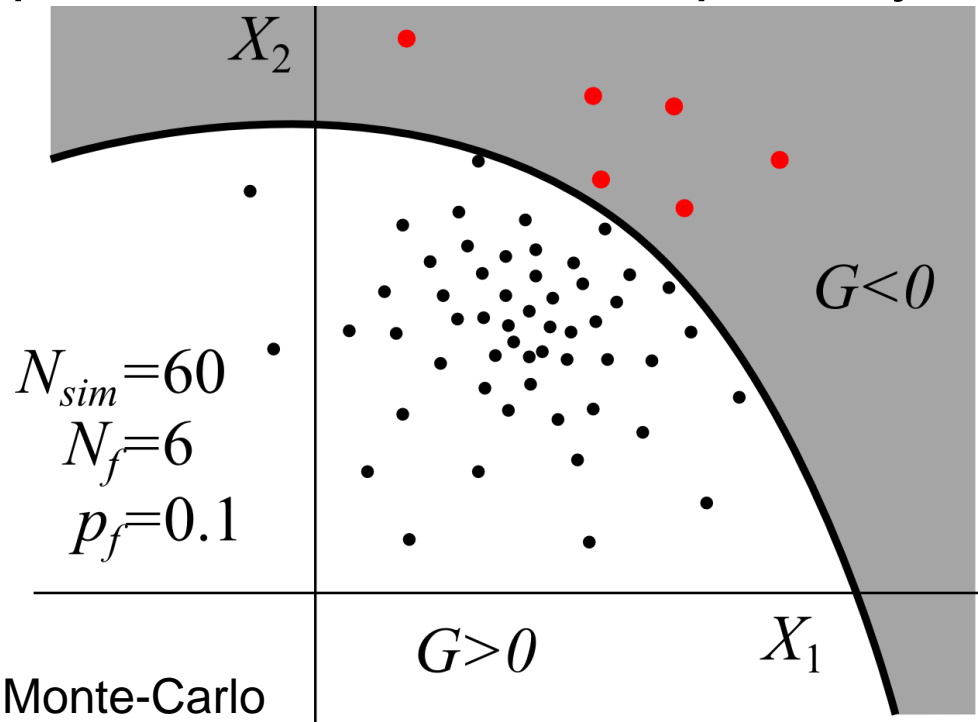
example taken from <http://fast10.vsb.cz/krejsa> (Martin Krejsa)

Obrázek 3. Monte-Carlo integrace

# MC v pravděpodobnostních výpočtech

- generuje náhodné vstupy podle daných **rozdělovacích funkcí**
- poté určí, která realizace spadá do bezpečné oblasti či oblasti poruchy
- pravděpodobnost poruchy je počet realizací v oblasti poruchy dělený celkovým počtem

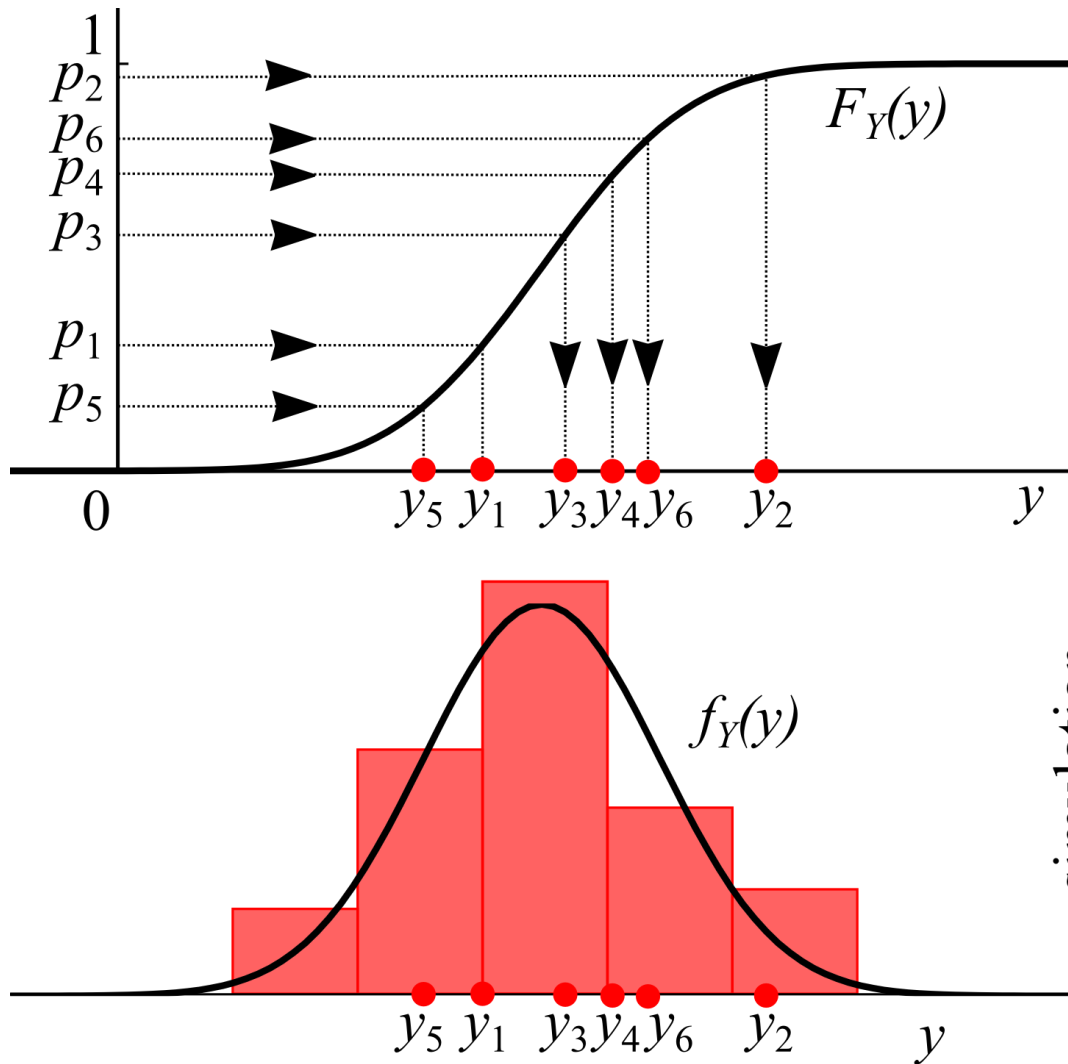
$$p_f \approx \frac{N_f}{N_{sim}}$$



Obrázek 4. Metoda Monte-Carlo



# Simulace náhodných veličin



$p_i$  generovány **náhodně** podle rovnoměrného rozdělení na intervalu 0-1

$$y_i = F_Y^{-1}(p_i)$$

|   | variable |       |       |     |
|---|----------|-------|-------|-----|
| 1 | $x_1$    | $y_1$ | $z_1$ | ... |
| 2 | $x_2$    | $y_2$ | $z_2$ | ... |
| 3 | $x_3$    | $y_3$ | $z_3$ | ... |
| 4 | $x_4$    | $y_4$ | $z_4$ | ... |
| 5 | $x_5$    | $y_5$ | $z_5$ | ... |
| 6 | $x_6$    | $y_6$ | $z_6$ | ... |
| ⋮ | ⋮        | ⋮     | ⋮     | ⋮   |

simulation

Obrázek 5. Simulace náhodné veličiny

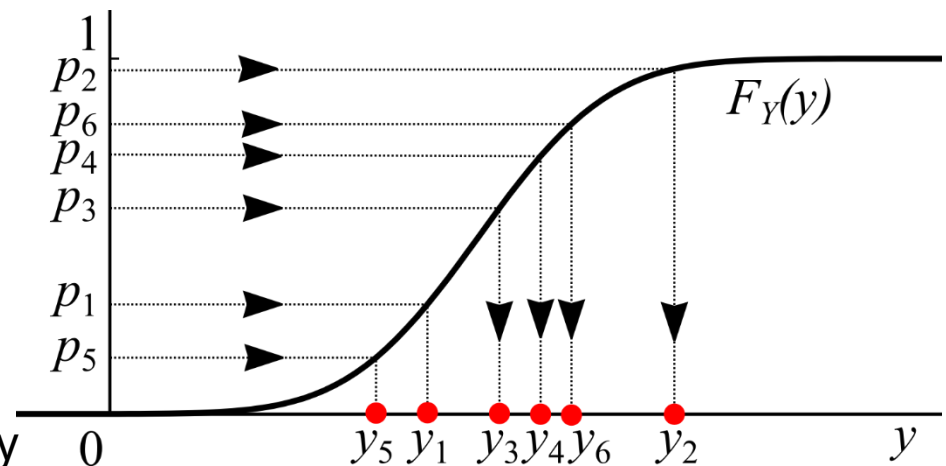
# Simulace náhodných veličin



- jednoduché na použití a porozumnění
- robustní a přesné
- dostupné dobré generátory náhodných čísel – statisticky nezávislé s dlouho periodou opakování
- „pseudonáhodná čísla“



- komplikované zavádění statistické závislosti mezi veličinami
- nutný velký počet opakování pro dosažení přesných odhadů výsledných statistik



Obrázek 6. Simulace náhodné veličiny





# Odhad pravděpodobnosti poruchy

–pro každý řádek simulační tabulky se vyčíslí rezerva spolehlivosti

– $N_f$  = počet negativních  $g(x_1, y_1, z_1, \dots)$

|            | variable |       |       |       | safety margin |                                 |
|------------|----------|-------|-------|-------|---------------|---------------------------------|
| simulation | 1        | $x_1$ | $y_1$ | $z_1$ | ...           | $g(x_1, y_1, z_1, \dots) < 0$ ✗ |
|            | 2        | $x_2$ | $y_2$ | $z_2$ | ...           | $g(x_1, y_1, z_1, \dots) > 0$ ✓ |
|            | 3        | $x_3$ | $y_3$ | $z_3$ | ...           | $g(x_1, y_1, z_1, \dots) > 0$ ✓ |
|            | 4        | $x_4$ | $y_4$ | $z_4$ | ...           | $g(x_1, y_1, z_1, \dots) > 0$ ✓ |
|            | 5        | $x_5$ | $y_5$ | $z_5$ | ...           | $g(x_1, y_1, z_1, \dots) < 0$ ✗ |
|            | 6        | $x_6$ | $y_6$ | $z_6$ | ...           | $g(x_1, y_1, z_1, \dots) > 0$ ✓ |
|            | ⋮        | ⋮     | ⋮     | ⋮     | ⋮             |                                 |

➔

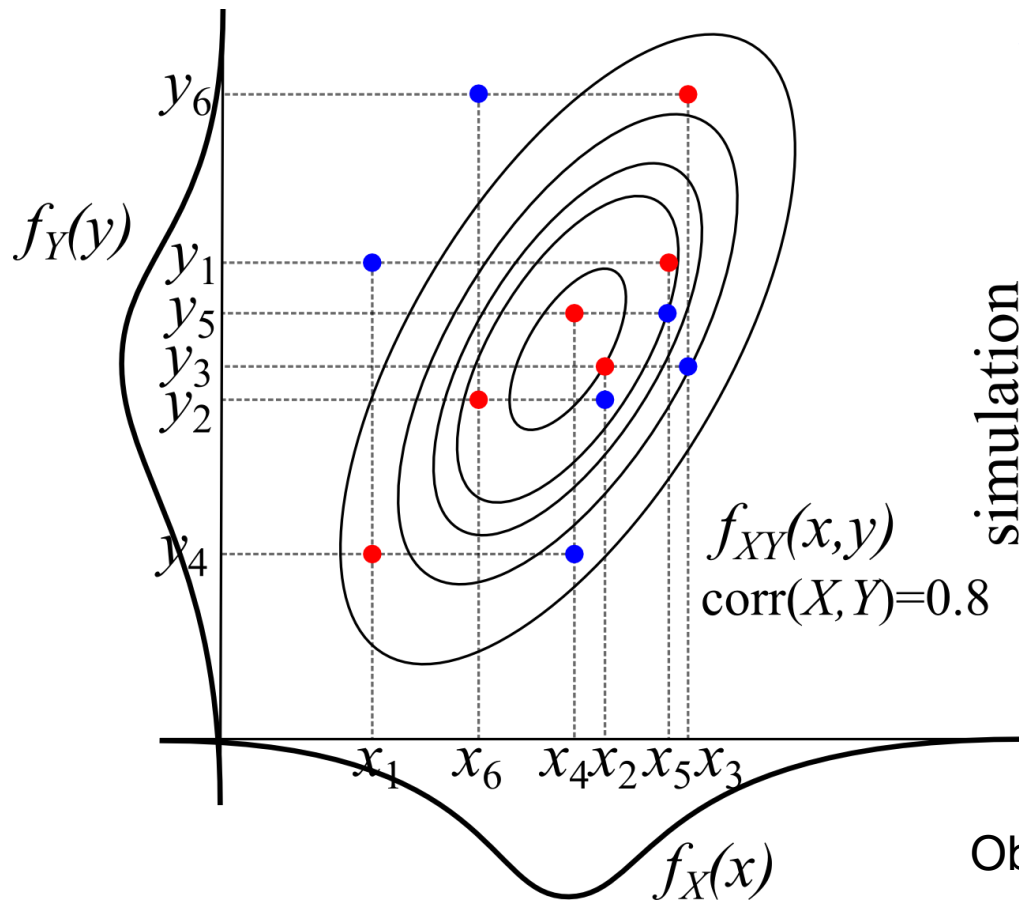
$$P_f \approx \frac{N_f}{N_{sim}}$$

Obrázek 7. Odhad pravděpodobnosti poruchy s použitím simulovaných vstupů



# Korelace náhodných veličin

– lze ji zavést **přehazováním pořadí** simulovaných realizací –  
toto bude podrobně probráno v části 6



$\text{corr}(X, Y) = 0.0$

$\text{corr}(X, Y) = 0.8$

| variable   |          |                 | variable   |          |                       |
|------------|----------|-----------------|------------|----------|-----------------------|
| simulation | 1        | $x_1 \quad y_1$ | simulation | 1        | $x_1 \rightarrow y_4$ |
|            | 2        | $x_2 \quad y_2$ |            | 2        | $x_2 \rightarrow y_3$ |
|            | 3        | $x_3 \quad y_3$ |            | 3        | $x_3 \rightarrow y_6$ |
|            | 4        | $x_4 \quad y_4$ |            | 4        | $x_4 \rightarrow y_5$ |
|            | 5        | $x_5 \quad y_5$ |            | 5        | $x_5 \rightarrow y_1$ |
|            | 6        | $x_6 \quad y_6$ |            | 6        | $x_6 \rightarrow y_2$ |
|            | $\vdots$ | $\vdots$        |            | $\vdots$ | $\vdots$              |

Obrázek 8. Zavádění korelace v MC

# Pravděpodobnost vyjímečných událostí

- Jaká je pravděpodobnost, že padne **panna**?
- Jaká je pravděpodobnost, že padne **orel**?
  
- jednoduše lze odhadnout hozením mince 100 krát a podělením počtu orlů a panen celkovým počtem hodů (100)
  
- Jaká je však pravděpodobnost, že mince zůstane **stát na hraně**?
  
- Kolik hodů potřebujeme?
- Selhání stavební konstrukce!



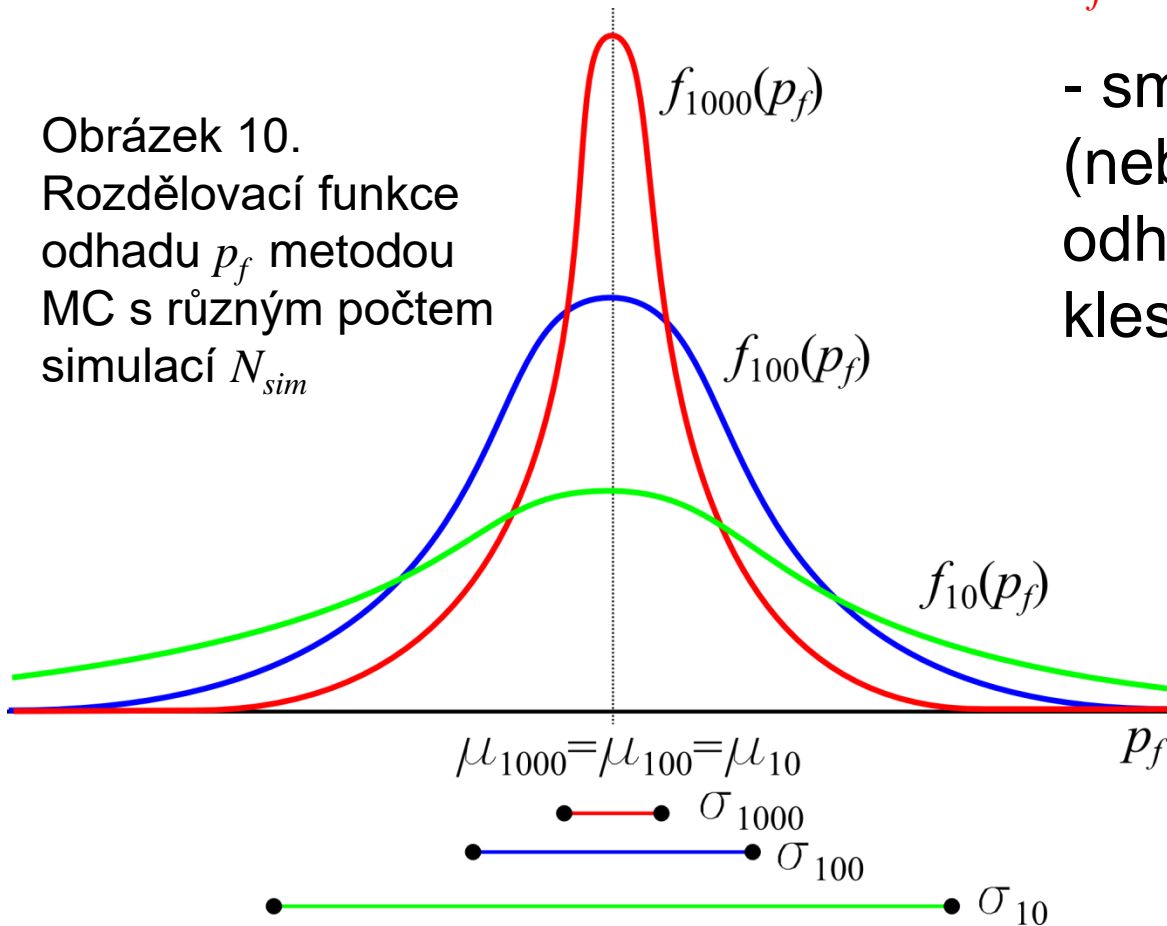
Obrázek 9. Jedno Euro



# Odhad chyby

–pokaždé se vygeneruje **jiná sada** pravděpodobností  $p_i \Rightarrow$  **jiné realizace** náhodných veličin  $\Rightarrow$  **jiná**  $p_f$

Obrázek 10.  
Rozdělovací funkce odhadu  $p_f$  metodou MC s různým počtem simulací  $N_{sim}$



- směrodatná odchylka (nebo **variční koeficient  $v$** ) odhadu pravděpodobnosti klesá s roustoucím  $N_{sim}$

$$v_{p_f} \approx \frac{1}{\sqrt{N_{sim} p_f}}$$



# Formální zápis metody MC

–definice „indikátorové“ funkce náhodného vektoru  $\mathbf{x}$

$$\mathbf{1}[g(\mathbf{x}) < 0] = \begin{cases} 0 & g(\mathbf{x}) \geq 0 \\ 1 & g(\mathbf{x}) < 0 \end{cases}$$

–1 značí že  $\mathbf{x}$  je v oblasti poruchy, 0 značí bezpečnou oblast

$$p_f = \int_{G < 0} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \int_{\Omega} \mathbf{1}[g(\mathbf{x}) < 0] f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$$

–numerické řešení integrálu

$$p_f = \frac{1}{N_{Sim}} \sum_{i=1}^{N_{Sim}} \mathbf{1}[g(\mathbf{x}_i) < 0] = \frac{N_f}{N_{Sim}}$$

tento člen je  
schován v  
generování  
náhodných  
veličin

# LHS - vylepšení vzorkování

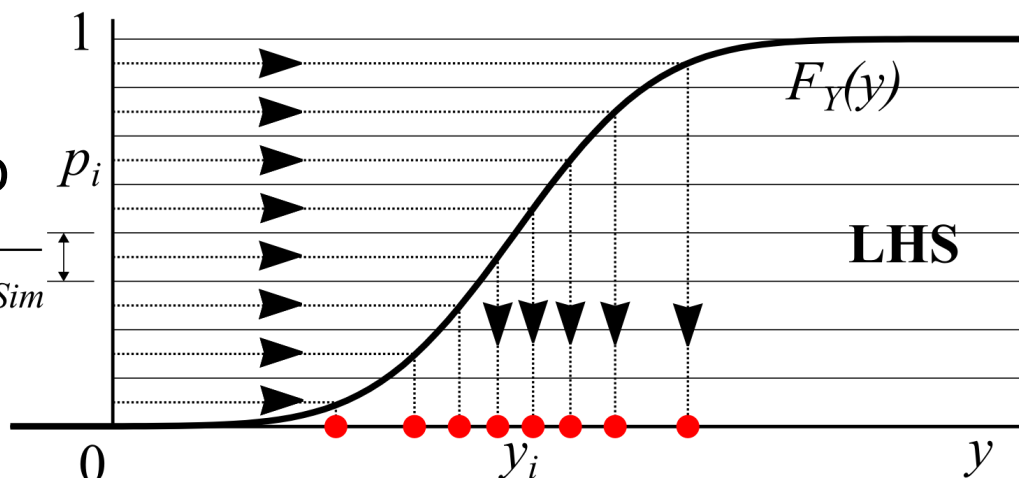
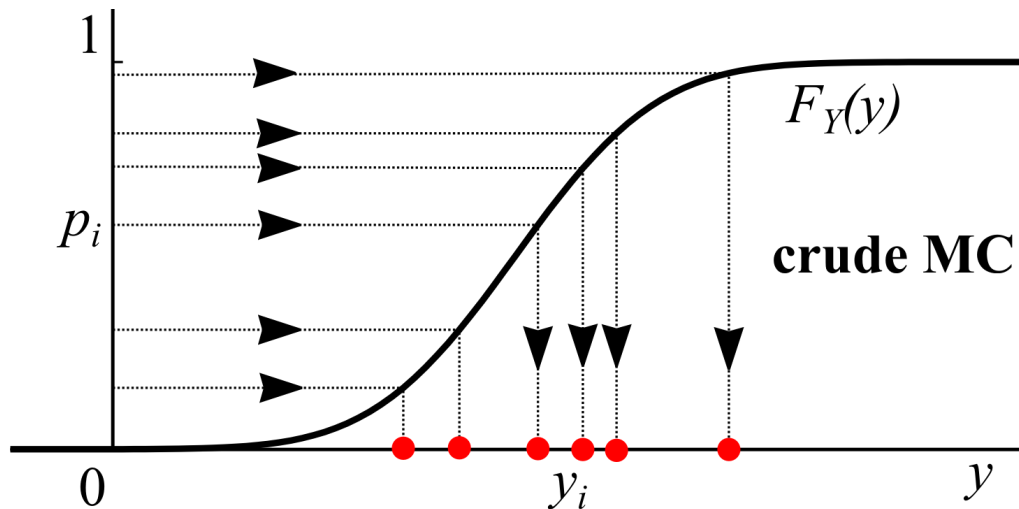
## MONTE CARLO

- $p_i$  vybráno náhodně
- jednoduché na použití i porozumnění
- možnost přidávat simulace

## LATIN HYPERCUBE

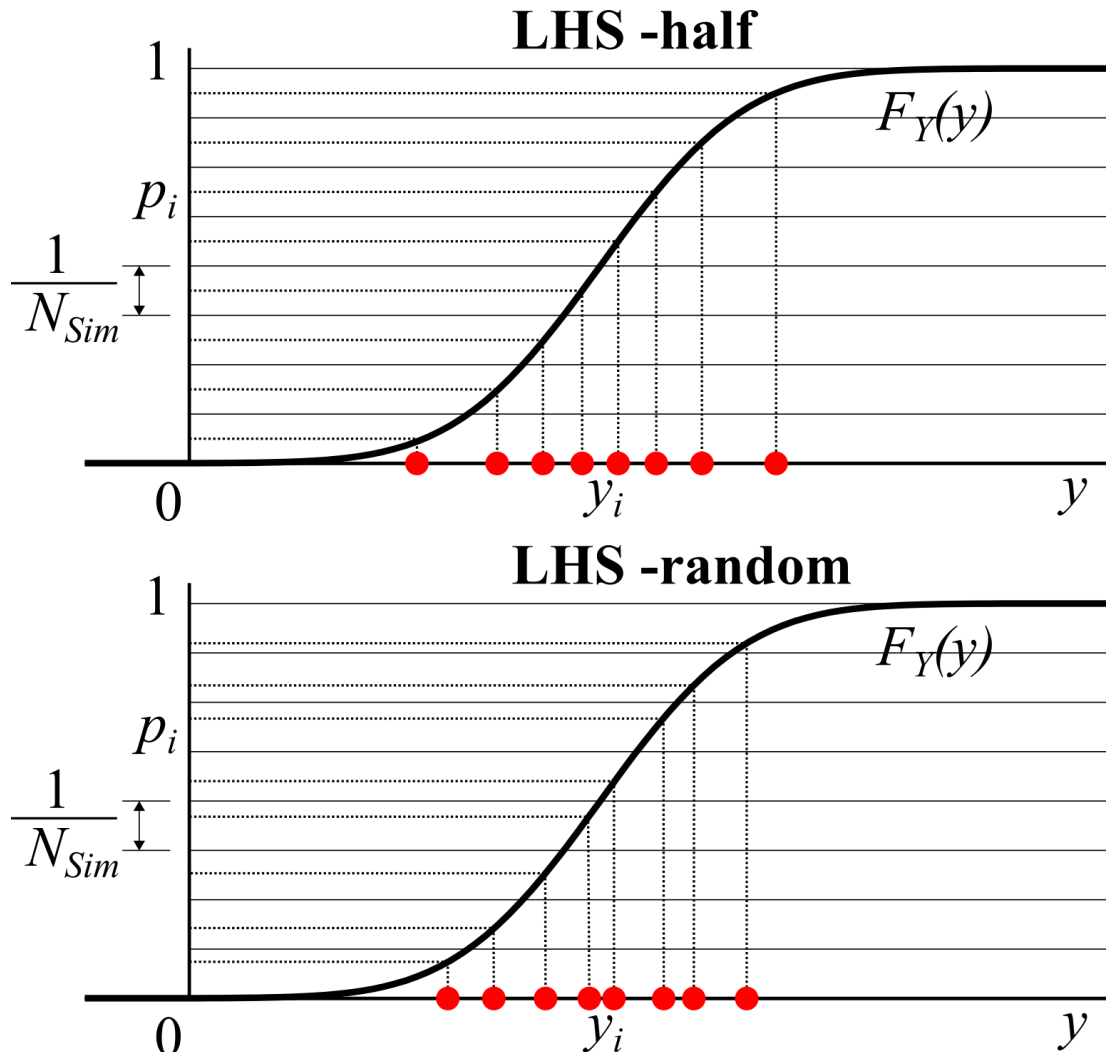
## SAMPLING

- jedno  $p_i$  vybráno z každého z  $N_{sim}$  intervalů
- vysoká přesnost pro nízká  $N_{sim}$
- složité přidávání simulací



Obrázek 11. MC & LHS

# Různé druhy LHS



Obrázek 12. LHS half & LHS random

## LHS – half (median)

–používá medián z každého intervalu

$$p_i = \frac{i - 0.5}{N_{Sim}}$$

pro  $i \in \{1, 2, 3, \dots, N_{Sim}\}$

## LHS - random

–používá náhodné  $p$  z každého intervalu

$$p_i = \frac{i - 1 + \text{rand}(0,1)}{N_{Sim}}$$

pro  $i \in \{1, 2, 3, \dots, N_{Sim}\}$



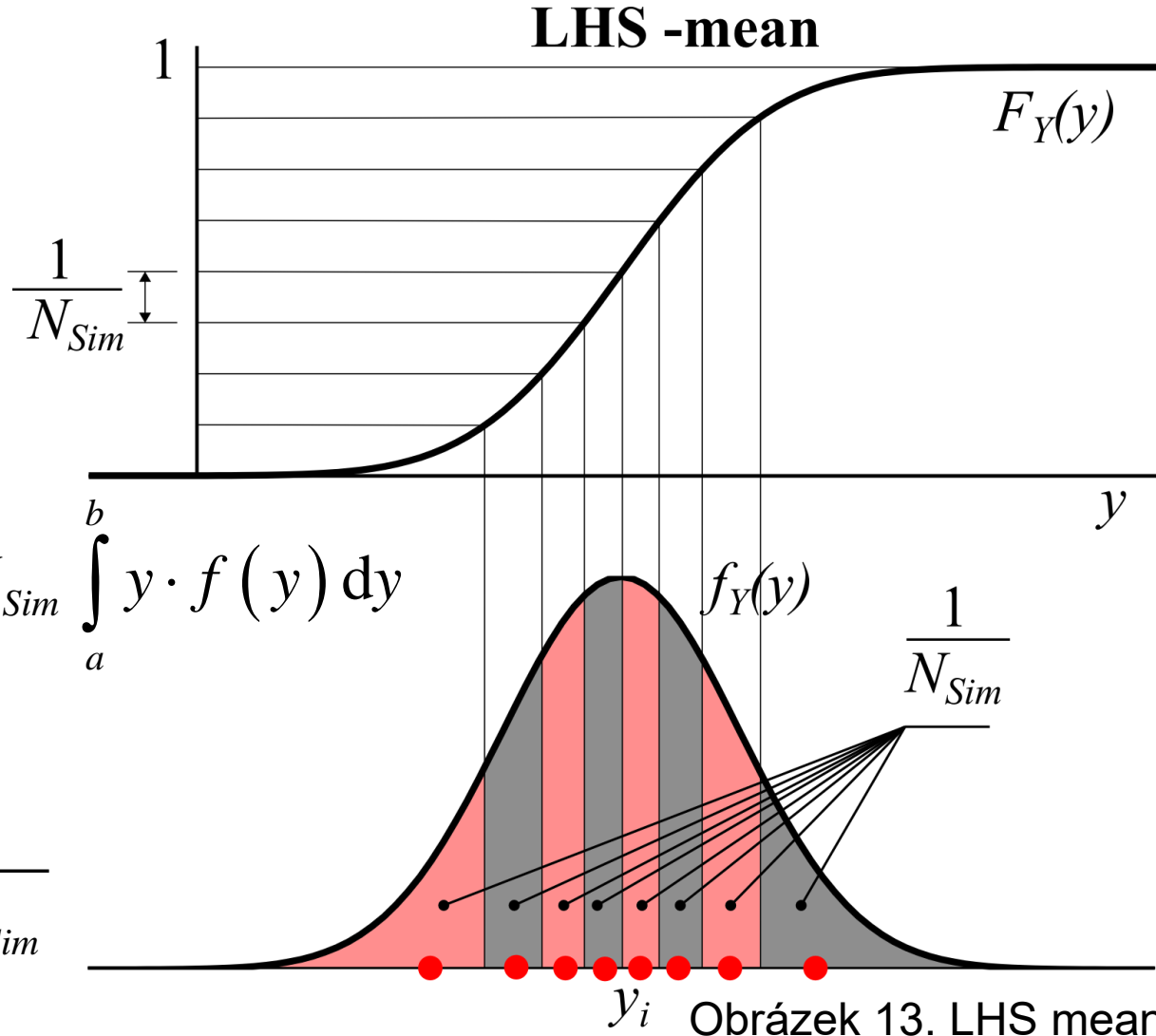
# Různé druhy LHS

## LHS - mean

–používá střední hodnotu každého z intervalů = **těžiště** intervalu

$$y_i = \frac{\int_a^b y \cdot f(y) dy}{\int_a^b f(y) dy} = N_{Sim} \int_a^b y \cdot f(y) dy$$

$$\text{kde } a = \frac{i-1}{N_{Sim}} \quad b = \frac{i}{N_{Sim}}$$



Obrázek 13. LHS mean



# Importance sampling

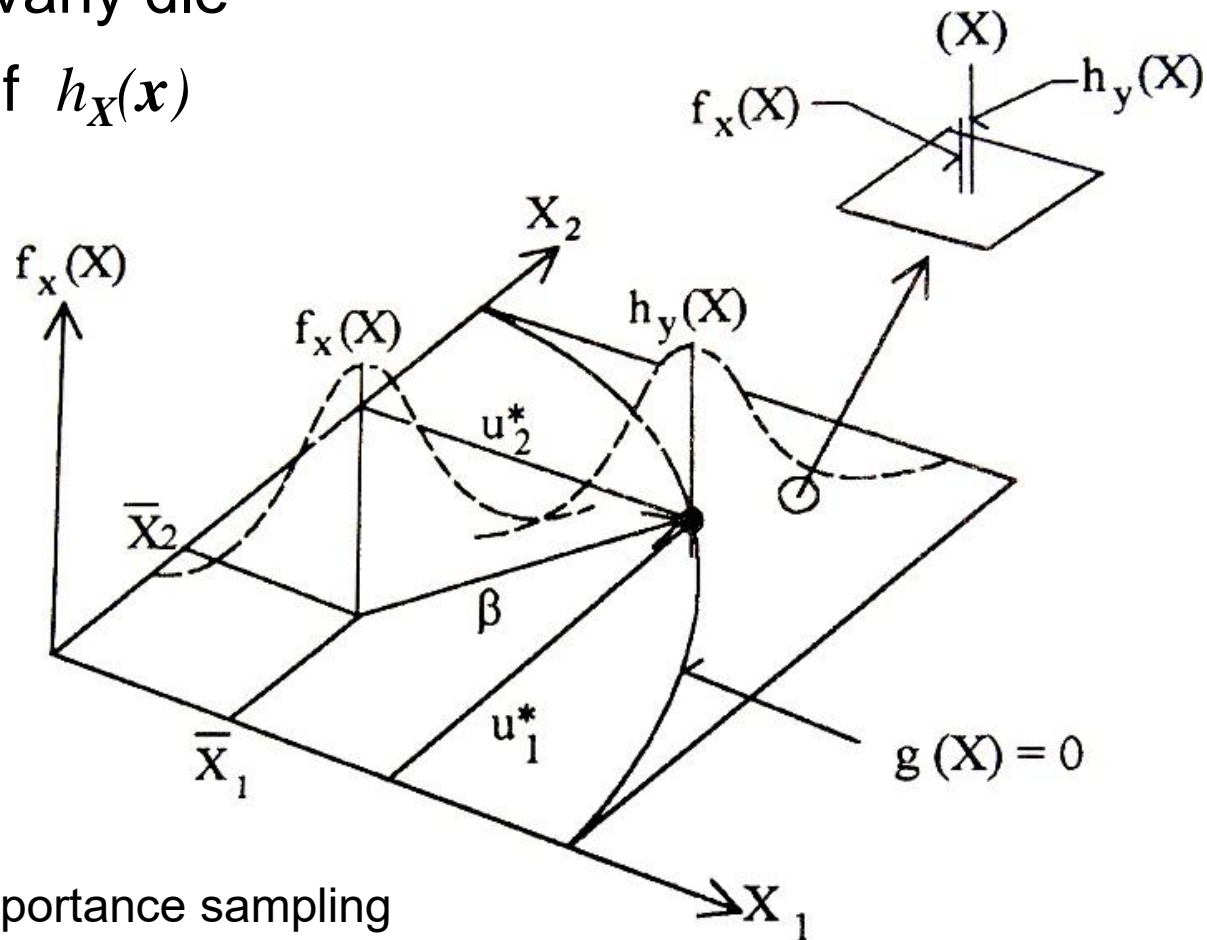
- základní myšlenka je generovat simulace **blíže oblasti poruchy**
- simulace jsou generovány dle

libovolně zvolené pdf  $h_X(\mathbf{x})$

**(váhová funkce)**

- pak je ale nutné  
výsledek opravit  
členem

$$\frac{f_X(\mathbf{x})}{h_X(\mathbf{x})}$$



Obrázek 14. Importance sampling



# Formální zápis Importance Sampling

–dříve uvedený formální zápis metody MC

$$p_f = \int_{\Omega} \mathbf{1}[g(\mathbf{x}) < 0] f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$$

–ted' však probíhá generování podle  $h_{\mathbf{X}}(\mathbf{x})$ , je tedy nutné rovnici opravit

$$p_f = \int_{\Omega} \mathbf{1}[g(\mathbf{x}) < 0] \frac{f_{\mathbf{X}}(\mathbf{x})}{h_{\mathbf{X}}(\mathbf{x})} h_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}$$

opravný člen

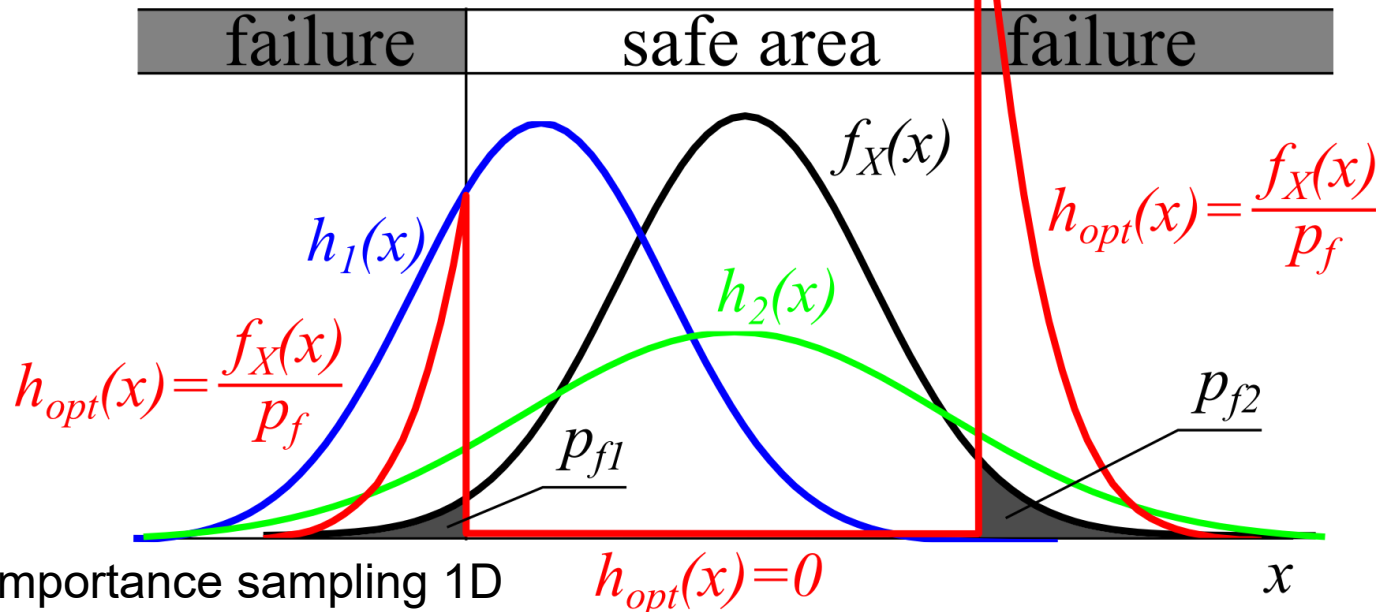
člen schovaný v generování náhodných veličin

–numerické řešení integrálu

$$p_f = \frac{1}{N_{Sim}} \sum_{i=1}^{N_{Sim}} \mathbf{1}[g(\mathbf{x}_i) < 0] \frac{f_{\mathbf{X}}(\mathbf{x}_i)}{h_{\mathbf{X}}(\mathbf{x}_i)}$$

# Která pdf $h_X(x)$ je vhodná

- obecné řešení neexistuje
- většinou je dobré jen zvýšit směrodatné odchylky rozdělovacích funkcí = **importance sampling okolo středních hodnot**
- optimální** pdf  $h_X(x)$  je nulová v bezpečné oblasti a proporcionalní k  $f_X(x)$  v oblasti poruchy



Obrázek 15. Importance sampling 1D

$$h_{opt}(x) = 0$$



# Importance sampling - komentáře

- váhová funkce: problém heuristického odhadu – typ, střední hodnota, rozptyl, ...
- malý rozptyl**: realizace jsou většinou v bezpečné oblasti jako u metody MC
- velmi velký rozptyl**: nestabilní výpočet, velký rozptyl odhadu pravděpodobnosti poruchy
- umožňuje odhad malých pravděpodobností poruchy při nízkém počtu simulací (např. 1000)
- daleko lepší ve srovnání s MC metodou při stejném počtu simulací:

$$D^2 [P_f]_{\text{Importance Sampling}} \leq D^2 [P_f]_{\text{Monte Carlo}}$$