SPECIAL MOBILITY STRAND

Reliability Analysis Techniques
Juan Sepúlveda
Novi Sad, Serbia, 22nd February 2020

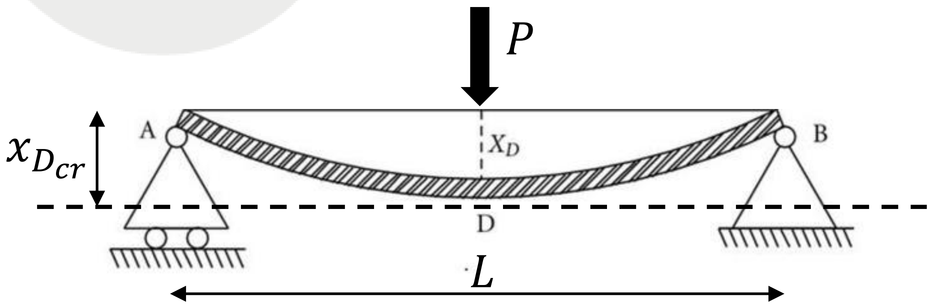
Juan Sepúlveda [1] Aalborg University

Co-funded by the Erasmus+ Programme of the European Union

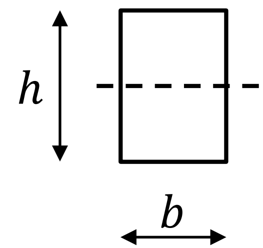
Presentation outline

- Intoduction
    - Deterministic design
    - Probabilistic design
    - Variable transformation
- First Order Reliability Method (FORM)
- Second Order Reliability Method (SORM)
- Crude Monte Carlo Simulation (CMCS)
- Importance Sampling (IS)
- Asymptotic Samplig (AS)

Co-funded by the Erasmus+ Programme of the European Union



Cross section:



$$I = \frac{bh^3}{12}$$

Known data:

$$h = 0.5$$
$$L = 15$$
$$E = 30 \cdot 10^9$$
$$x_{D_{cr}} = 0.03$$
$$P = 100{,}000$$

Limit state function

$$g = 1 - \frac{x_D}{x_{D_{cr}}}$$

Failure if: $g < 0$

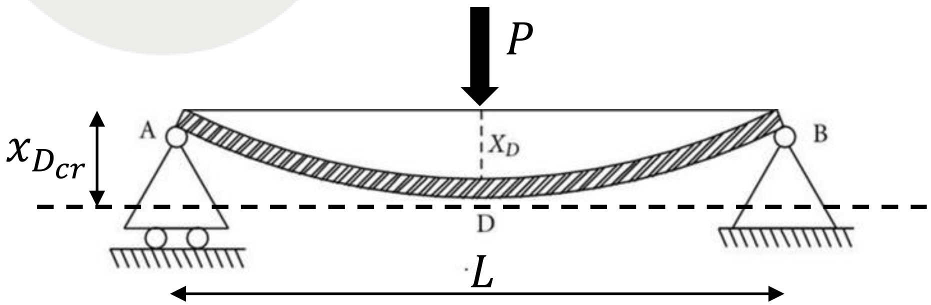
$b = ?$

where

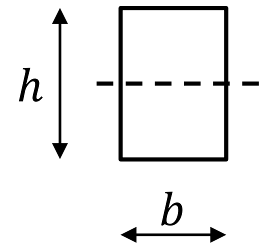
$$x_D = \frac{P}{K_D} \quad \text{and} \quad K_D = \frac{48EI}{L^3}$$

then

$$g = 1 - \frac{PL^3}{4Ebh^3 x_{D_{cr}}} > 0 \quad \Rightarrow \quad b > \frac{PL^3}{4Eh^3 x_{D_{cr}}} = 0.22$$



Cross section:



$$I = \frac{bh^3}{12}$$

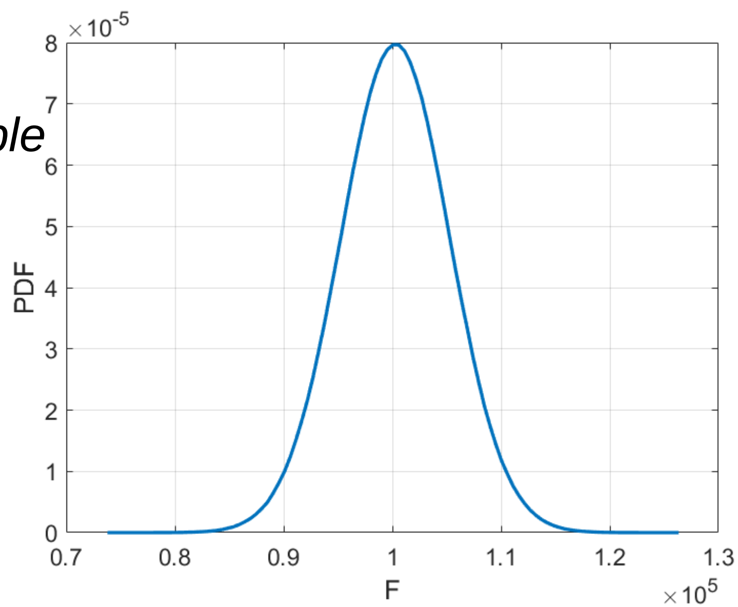
Uncertain loading:
P is a random variable

$$P \sim N(\mu_P, \sigma_P)$$

$$\mu_P = 100{,}000$$

$$\sigma_P = 5{,}000$$

$$C.o.V. = \frac{\sigma_P}{\mu_P} = 5\%$$



It is always convenient to use standard normal random variables, i.e.:

$$P = z\sigma_P + \mu_P$$

$$z \sim N(0,1)$$

A simple transformation from X_i to Z_i can be defined by the identity

$$\Phi(Z_i) = F_{X_i}(X_i)$$

where F_{X_i} is the distribution function for X_i. Given a realization $\boldsymbol{z}$ of $\boldsymbol{Z}$ a realization $\boldsymbol{x}$ of $\boldsymbol{X}$ can be determined by

$$x_i = F_{X_i}^{-1}(\Phi(z_i))$$

The failure probability is given as

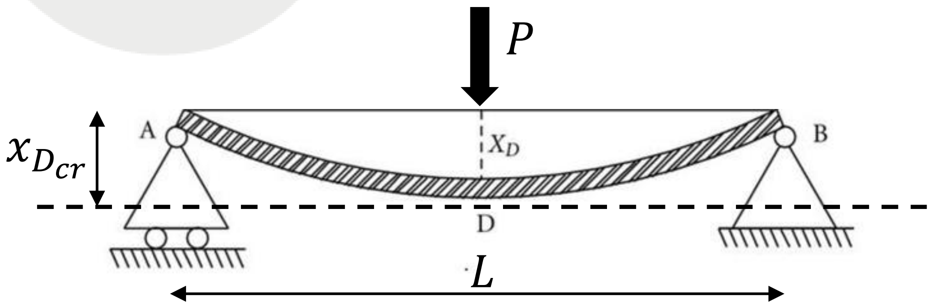
$$P_F = P(g < 0) = \int_F q(\boldsymbol{z})\mathrm{d}\boldsymbol{z}$$

Where $q(\boldsymbol{z})$ is the joint Probability Density Function (PDF) of the random variables and

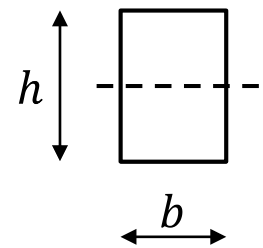
$$F = \{\boldsymbol{z} : g < 0\}$$

Denotes the "failure domain" or "failure region" which is a subset in the parameter space of $\boldsymbol{z}$

Co-funded by the
Erasmus+ Programme
of the European Union



Cross section:

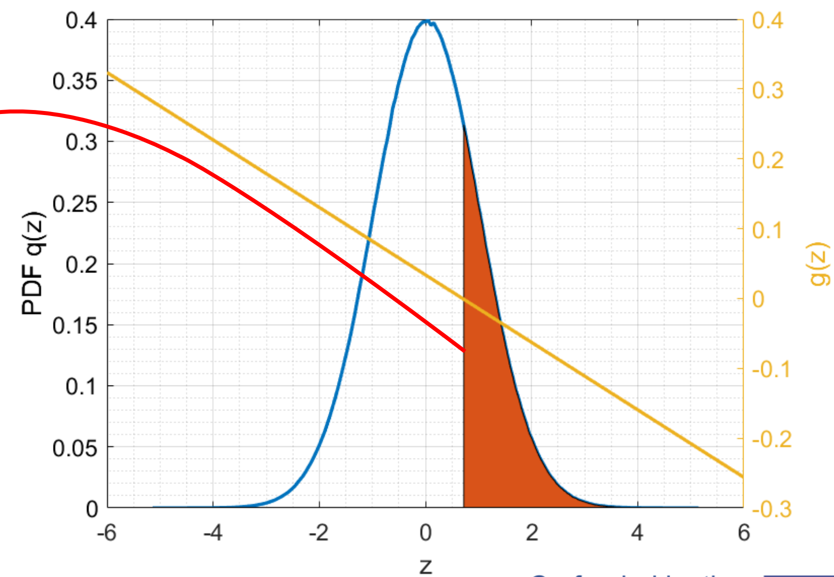


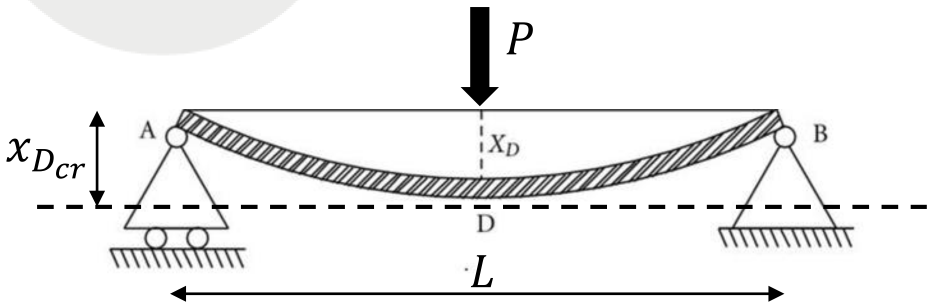
$$I = \frac{bh^3}{12}$$

In our example, using:

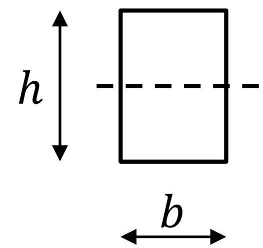
$$P_F = P(g < 0) = \int_F q(z)\mathrm{d}z$$

For $b = 0.23$ → $P_F = 0.242$





Cross section:



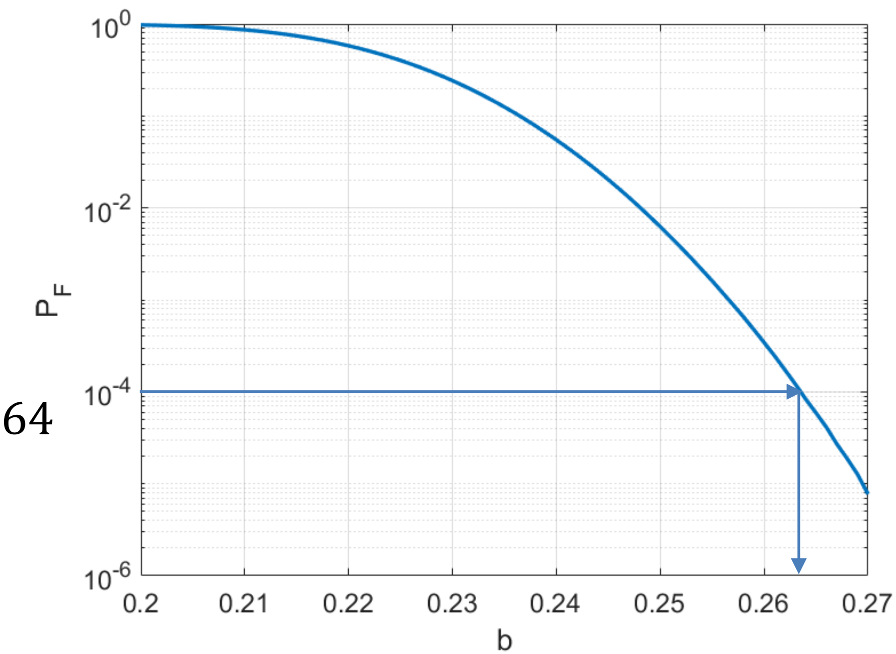
$$I = \frac{bh^3}{12}$$

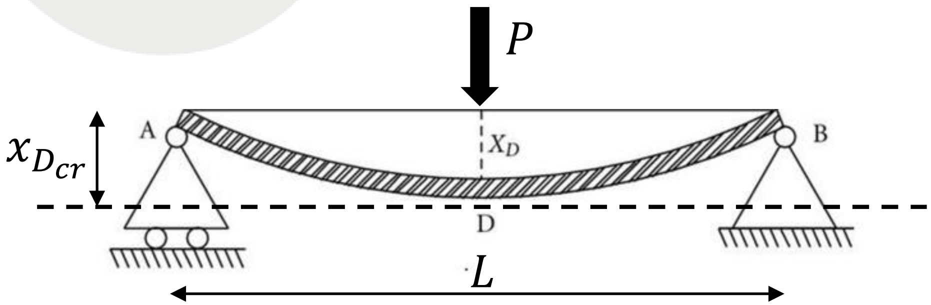
The design requirement now is the target reliability of the system

$$R^* = 1 - P_F^*$$

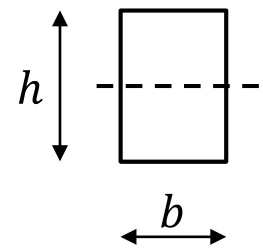
Or equivalently the target failure probability P_F^*

For example: $P_F^* = 10^{-4}$ ⇒ $b = 0.264$





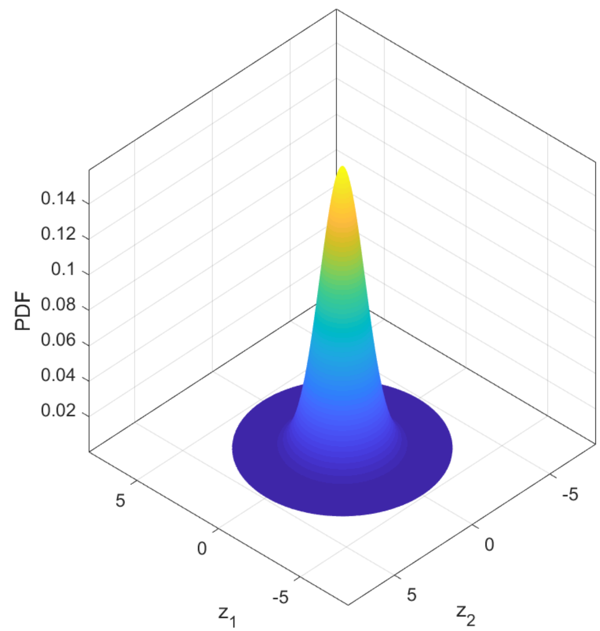
Cross section:



$$I = \frac{bh^3}{12}$$

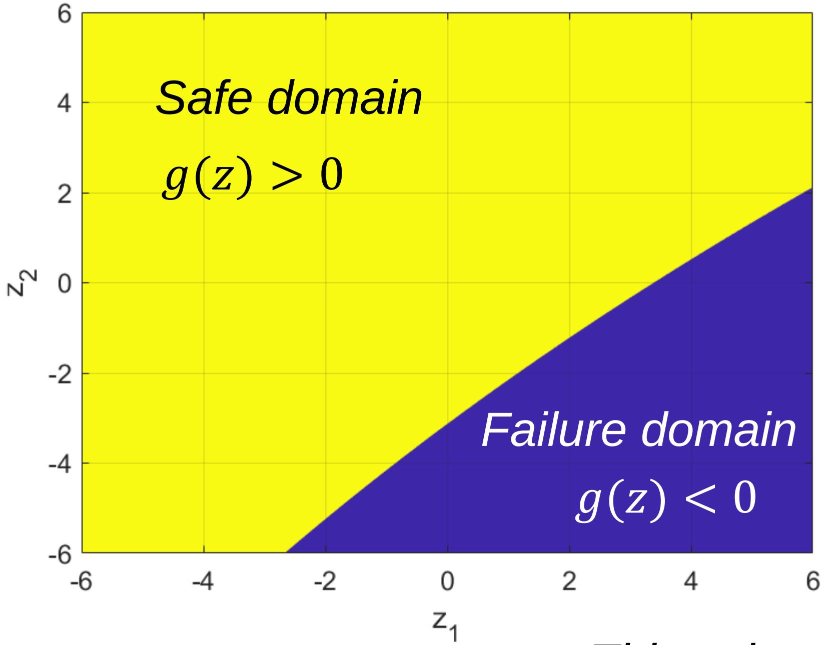
Probabilistic Model

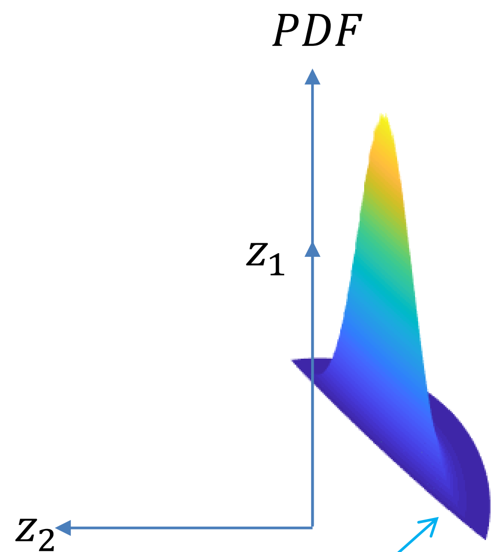
Random variable	Distribution	Mean value	C.o.V
P	Normal	100,000	5%
E	Normal	$30 \cdot 10^9$	2.5%



Co-funded by the Erasmus+ Programme of the European Union

For a given value of $b = 0.26$





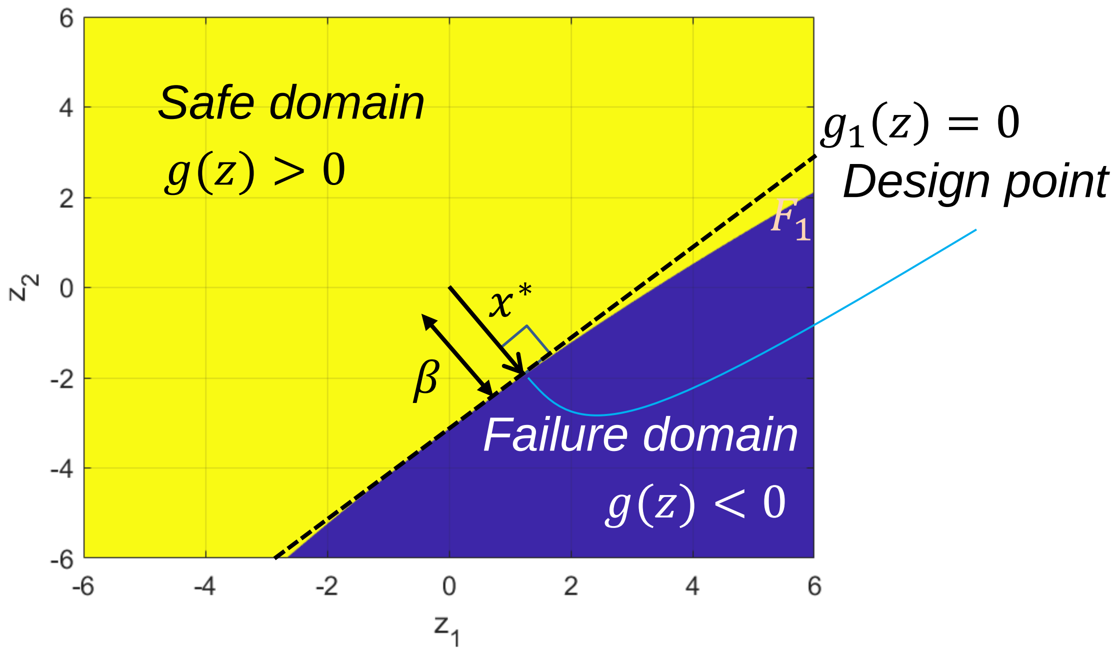
This volume represents the failure probability P_F

How to calculate it?

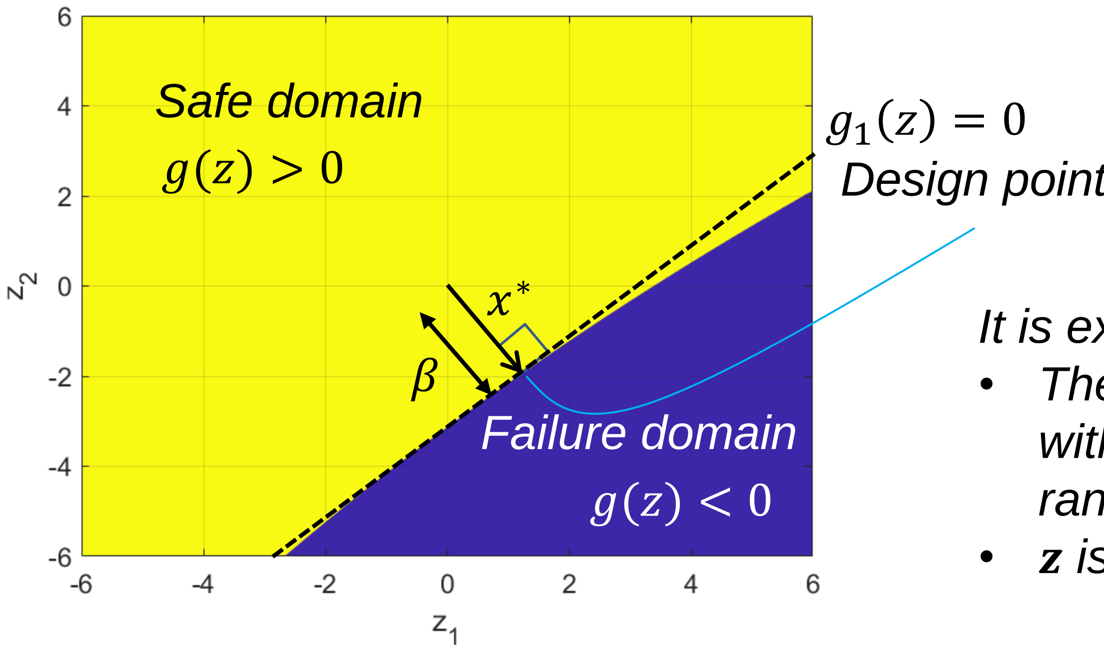
The failure domain is approximated by a linear half-space F_1.

The "design point" is the point in F_1 with minimal distance to the origin, i.e. with minimal Euclidean norm.

The maximum of the PDF in F_1 is at the design point.



Co-funded by the Erasmus+ Programme of the European Union



The approximation is given by

$$P_F \approx \widehat{P_F} = \Phi(-\beta)$$

It is exact when:

- *The limit state function is linear with respect to the vector of random variables* $\boldsymbol{z}$
- $\boldsymbol{z}$ *is Gaussian distributed*

For $b = 0.26$*, the approximation gives*

$$P_F \approx \widehat{P_F} = \Phi(-2.243) = 0.0124$$

Co-funded by the Erasmus+ Programme of the European Union

How to determinate β?

1. Guess $\boldsymbol{z}^{(0)}$. Set $i = 0$.

2. Calculate $g\left(\boldsymbol{z}^{(i)}\right)$ and $\nabla g\left(\boldsymbol{z}^{(i)}\right)$

3. Calculate an improved guess of β with

$$\boldsymbol{z}^{(i+1)} = \nabla g\left(\boldsymbol{z}^{(i)}\right) \frac{\nabla g\left(\boldsymbol{z}^{(i)}\right)^{T} \boldsymbol{z}^{(i)} - g(\boldsymbol{z}^{(i)})}{\nabla g(\boldsymbol{z}^{(i)})^{T} \nabla g(\boldsymbol{z}^{(i)})}$$

4. Calculate

$$\beta^{(i+1)} = \sqrt{(\boldsymbol{z}^{(i+1)})^{T} \boldsymbol{z}^{(i+1)}}$$

5. If $|\beta^{(i+1)} - \beta^{(i)}| < 10^{-3}$ then stop. Else $i = i + 1$ and go to step 2.

Co-funded by the
Erasmus+ Programme
of the European Union

+

- *The estimator does not depend of the order of* $\widehat{P_F}$
- *Under particular conditions leads to the exact solution*

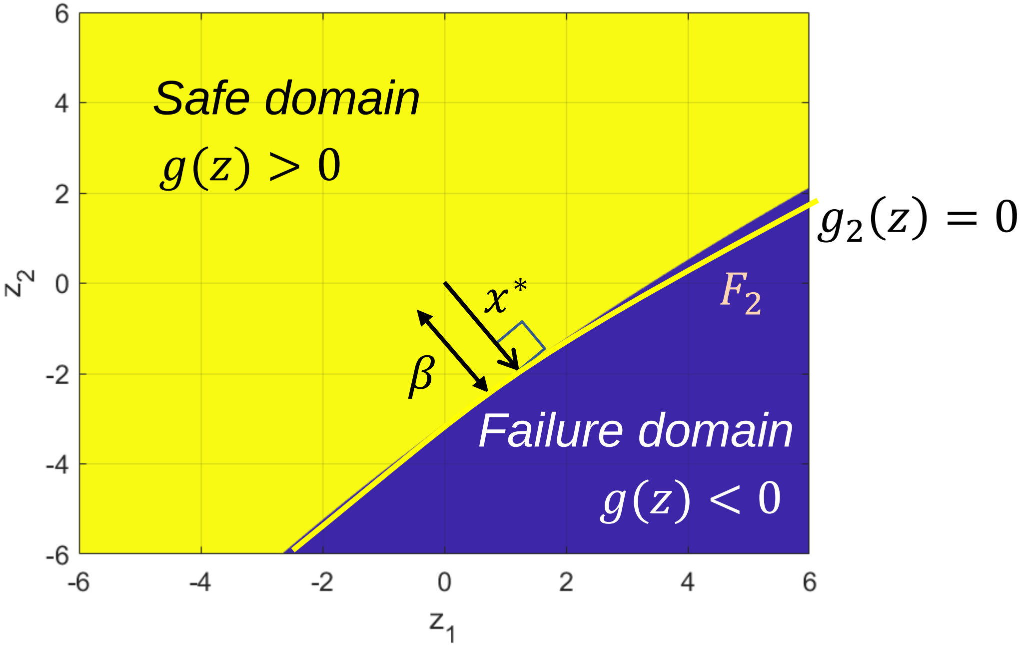
−

- *The linearization introduce an error*
- *It is found that is not accurate in high dimensions*
- *It is necessary to find the design point*
  - *It might be expensive in high dimensions*

Co-funded by the Erasmus+ Programme of the European Union

It intends to improve FORM taking into account the curvature of limit state function at the design point.

The failure boundary is approximated by an hyper-paraboloid with limit state function g_2.



Co-funded by the Erasmus+ Programme of the European Union

The SORM approximation is given by

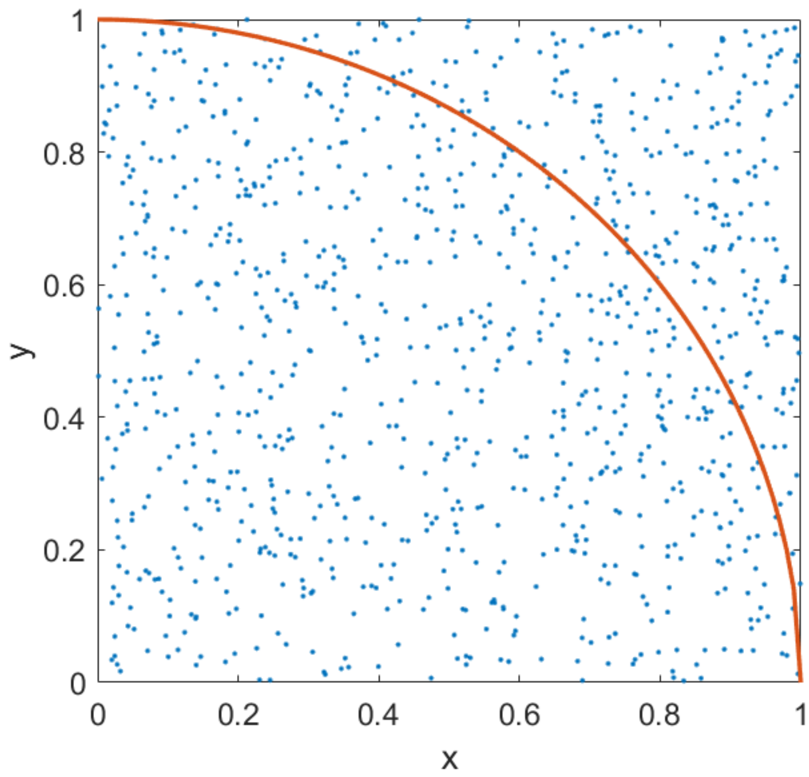
$$P_F \approx \widehat{P_F} = \Phi(-\beta)\prod_{i=2}^{n}(1 + c_i\beta)^{-\frac{1}{2}}, \qquad \beta \to \infty$$

Where c_i, $i = 2, \ldots, n$ *are the principal curvatures of the paraboloid at the design point.*

It is also assumed that $-c_i < 1/\beta$ *for all* $i = 2, \ldots, n$

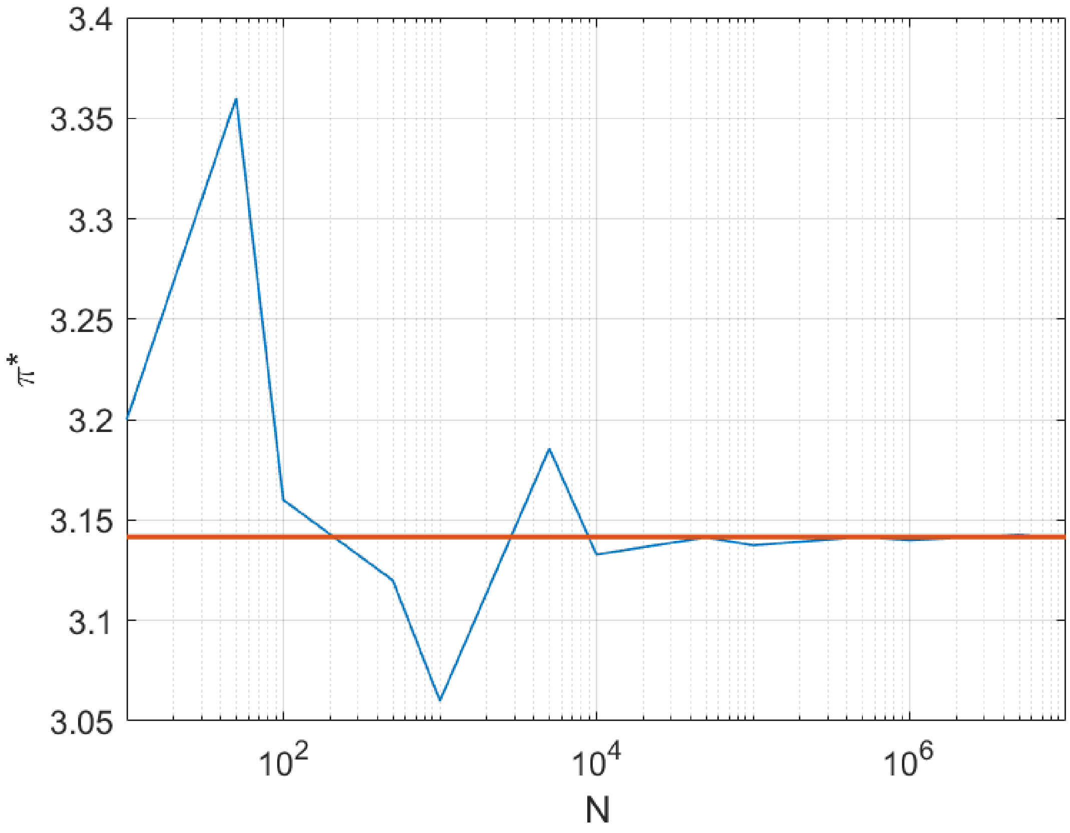
How to calculate the area of a circle? And implicitly π?

- *Simulate N random numbers into a unit square with area $A = 1$*
- *Draw a ¼ of a unit-radio circle with area $A_1 = \pi/4$*
- *Count the points inside the circle N_1*
- *The ratio $\frac{A_1}{A} = \frac{N_1}{N}$*
- *Then $\pi \approx 4\frac{N_1}{N}$*



Co-funded by the Erasmus+ Programme of the European Union

How is affected the estimation of $\pi \approx \pi^*$ by the number of simulations?

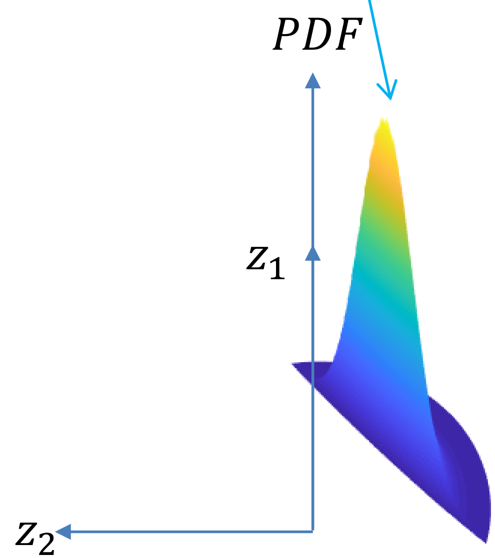


Funded by the
+ Programme
European Union

We can use the same approach to calculate this volume

- *A large number N of realizations $\hat{\boldsymbol{z}}$ of basic random variables is simulated using $q(\boldsymbol{z})$*

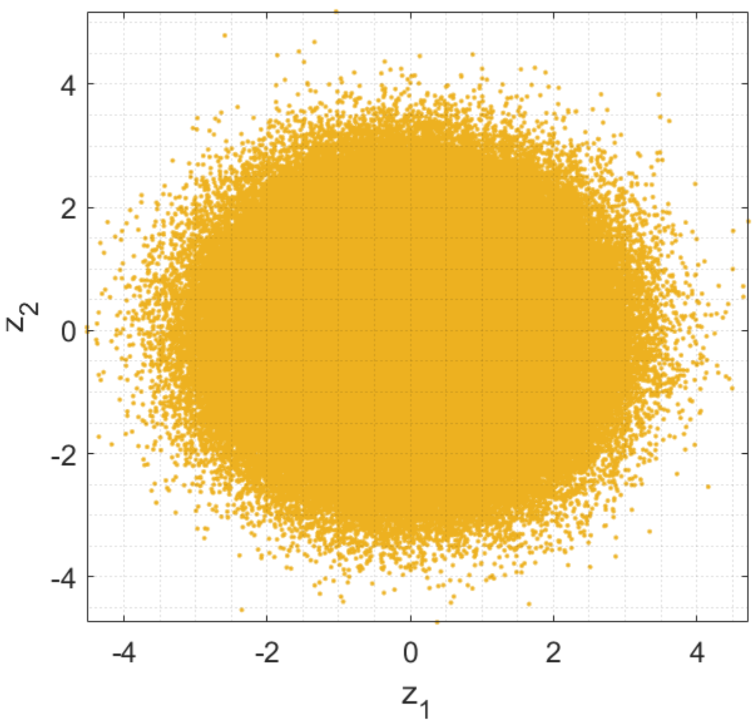
- *The failure probability estimator $\widehat{P_F}$ is obtained by counting the number of realizations in the failure domain*



Co-funded by the Erasmus+ Programme of the European Union

Procedure

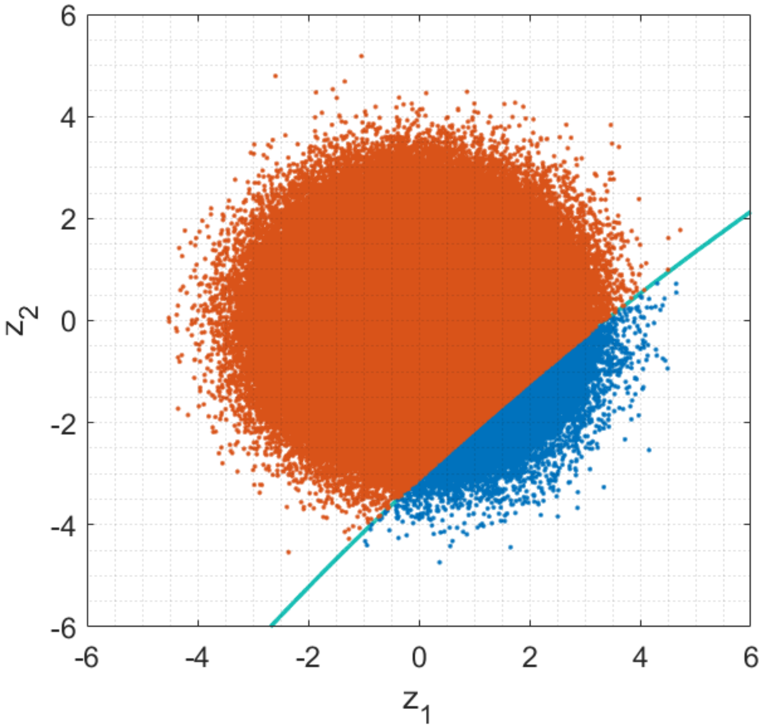
- N *realizations of standard normal random variables are simulated*
  - *This is the only step where the information regarding the distribution of the random variables in included*



Procedure

1. N realizations of standard normal random variables are simulated
   - This is the only step where the information regarding the distribution of the random variables is included

2. Identify the number N_1 of realizations in the failure domain
   - This is the only step where the information regarding the limit state function is included



Co-funded by the Erasmus+ Programme of the European Union

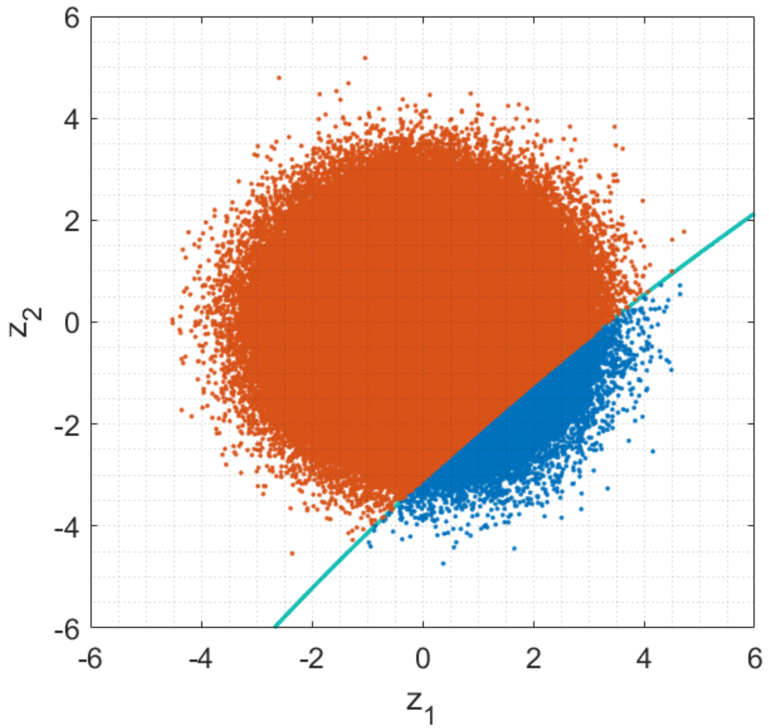
Procedure

1. *N realizations of standard normal random variables are simulated*
   - *This is the only step where the information regarding the distribution of the random variables is included*

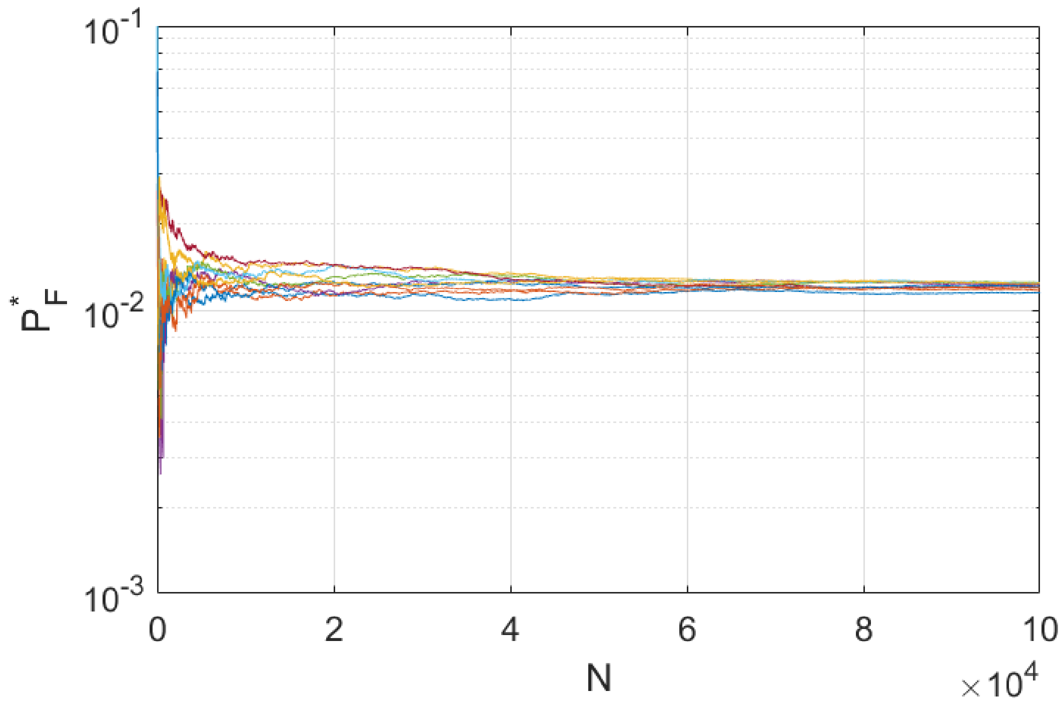
2. *Identify the number N_1 of realizations in the failure domain*
   - *This is the only step where the information regarding the limit state function is included*

3. *The estimator is given by*

$$\widehat{P_F} = \frac{N_1}{N}$$



In our example, using $b = 0.26$, how N affects the variability of the estimator $\widehat{P_F}$?



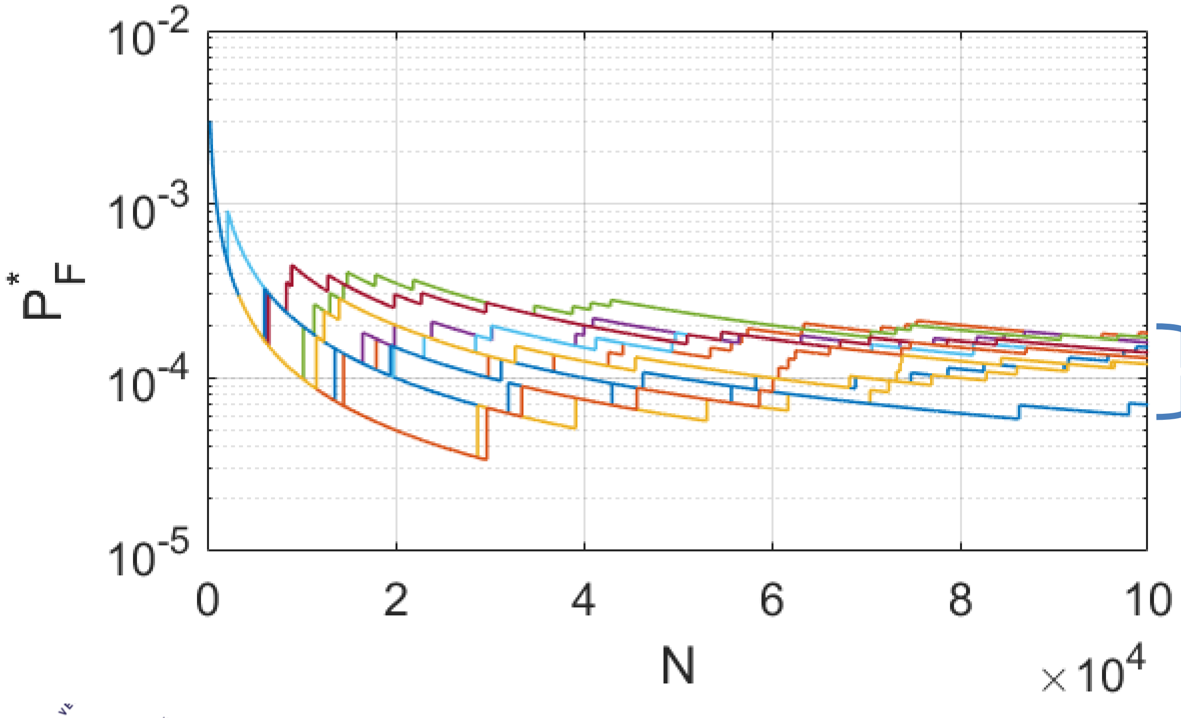
$$C.o.V(\widehat{P_F}) = \sqrt{\frac{1-\widehat{P_F}}{N}}$$

If $\widehat{P_F} = 0.0124$ *then*

N	$C.o.V(\widehat{P_F})$
10	282%
100	89%
1,000	28%
10,000	9%
100,000	3%

Co-funded by the Erasmus+ Programme of the European Union

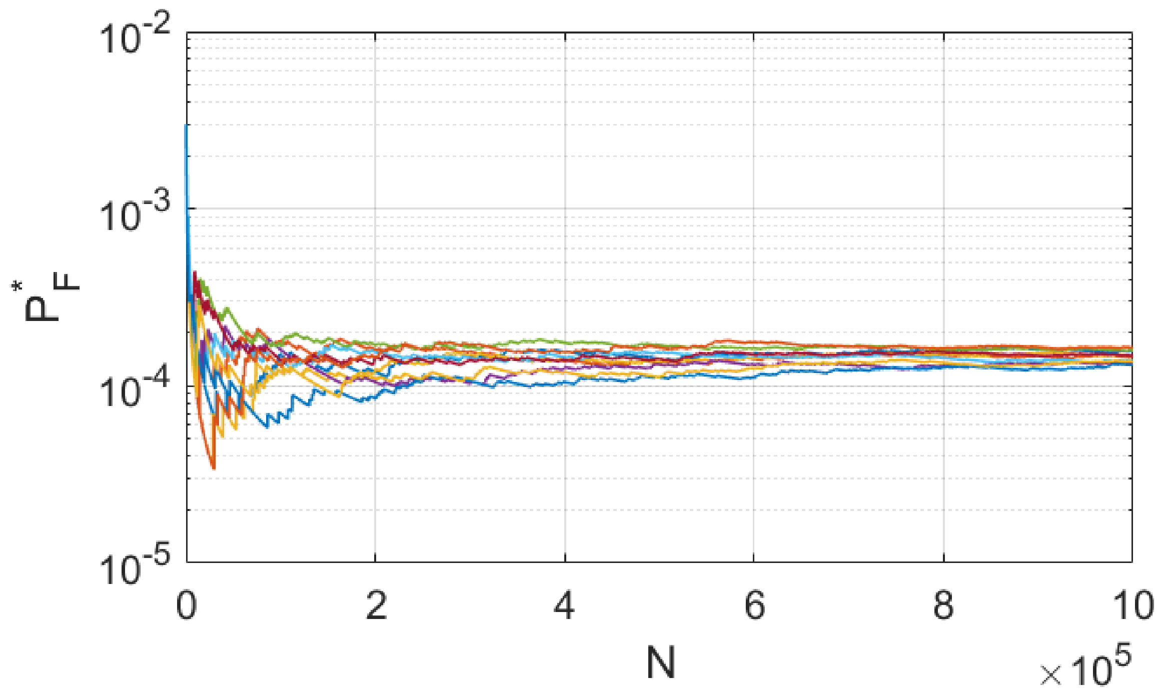
In our example, using $b = 0.285$*, i.e.* $\widehat{P_F} = 1.5 \times 10^{-4}$*. How* $\widehat{P_F}$ *affects* $C.o.V(\widehat{P_F})$*?*



It seems that there is not enough accuracy

Co-funded by the Erasmus+ Programme of the European Union

In our example, using $b = 0.285$, *i.e.* $\widehat{P_F} = 1.5 \times 10^{-4}$. *How* $\widehat{P_F}$ *affects* $C.o.V(\widehat{P_F})$?



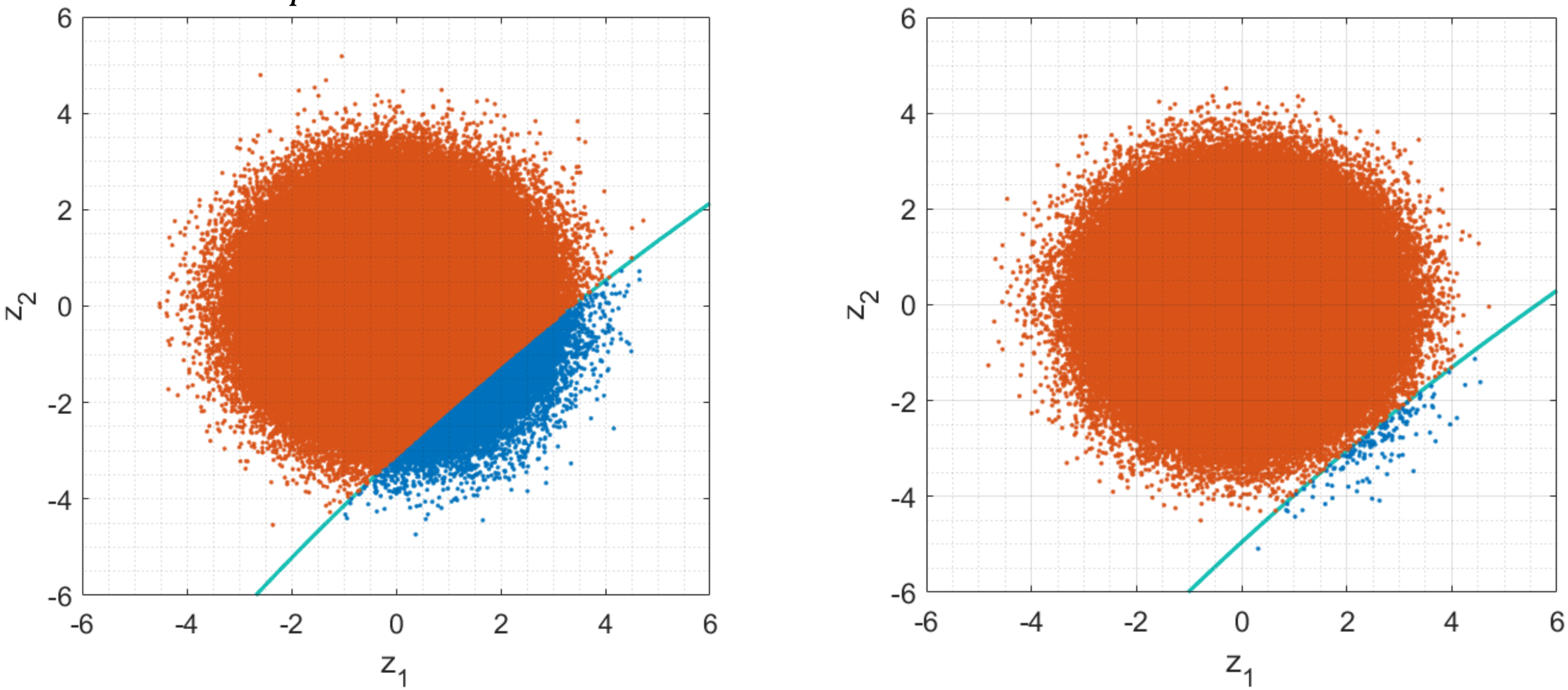
N	$C.o.V(\widehat{P_F})$
10	2581%
100	816%
1,000	258%
10,000	82%
100,000	26%
1,000,000	8% ✓

$C.o.V(\widehat{P_F}) < 10\%$ is normally accepted

Co-funded by the Erasmus+ Programme of the European Union

$b = 0.260$
$\widehat{P_F} = 1.24 \times 10^{-2}$

$b = 0.285$
$\widehat{P_F} = 1.5 \times 10^{-4}$



When the limit state function is far from the origin is more difficult to obtain realizations in the failure domain

Co-funded by the Erasmus+ Programme of the European Union

+

- *Independent of the number of random variables*
- *Independent of the shape of the limit state function*
- *The estimator is unbiased*

−

- *The number of samples required is proportional to* $1/P_F$

Co-funded by the Erasmus+ Programme of the European Union

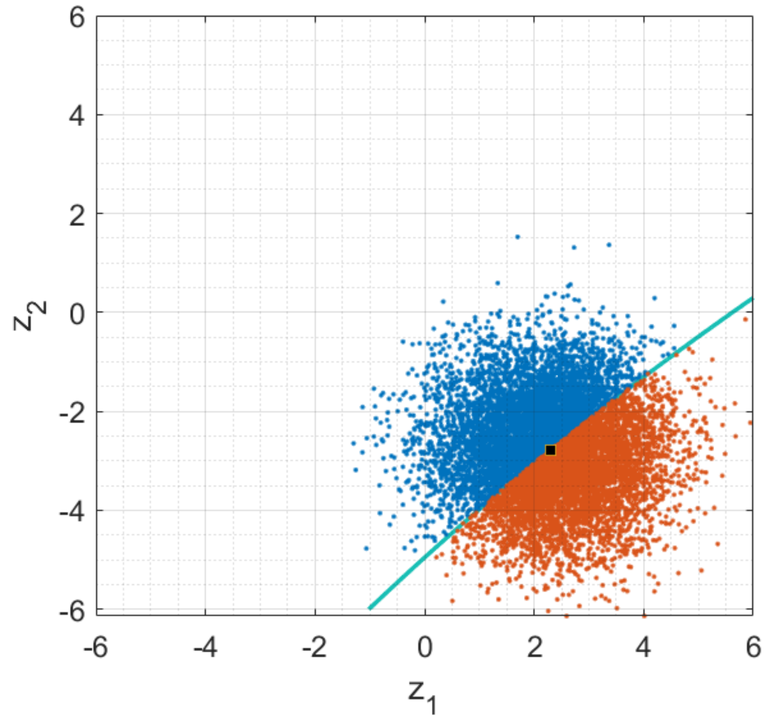
The idea is to concentrate the realizations in an area with large contribution to $\widehat{P_F}$

$$P_F = \int_{g<0} q(\mathbf{z})\mathrm{d}\mathbf{z} = \int_{g<0} \frac{q(\mathbf{z})}{p(\mathbf{z})} p(\mathbf{z})\mathrm{d}\mathbf{z}$$

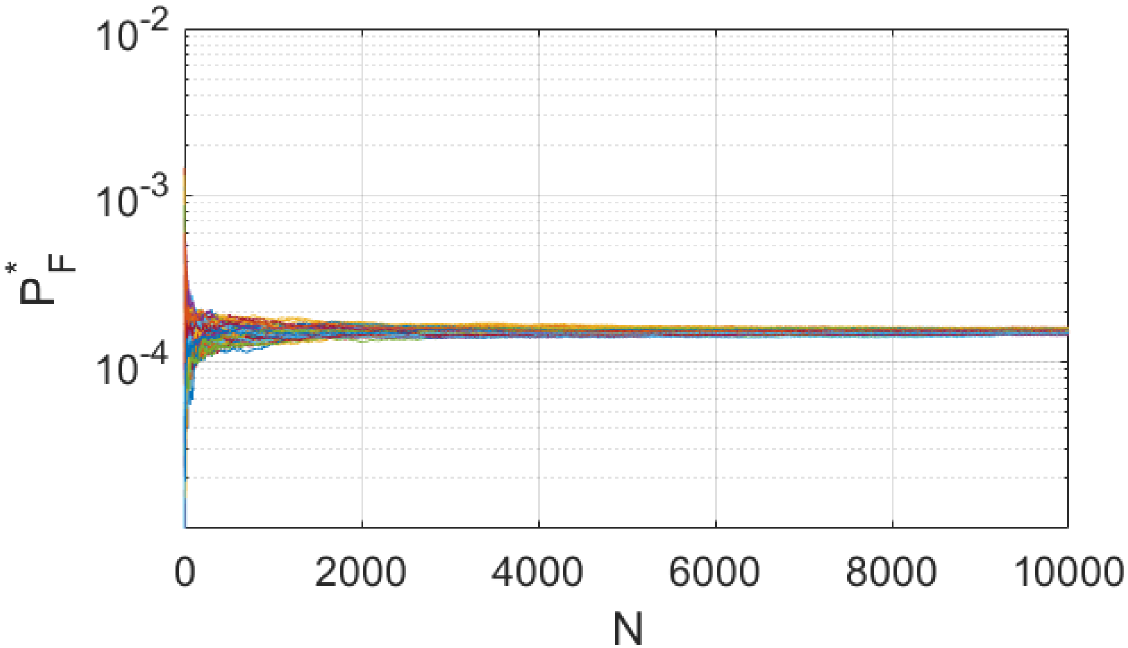
Then, the estimator is given by

$$\widehat{P_F} = \frac{1}{N}\sum_{k=1}^{N} \frac{q(\mathbf{z}'_k)}{p(\mathbf{z}'_k)} I(\mathbf{z}'_k \in F)$$

Where $\{\mathbf{z}'_k : k = 1, \dots, N\}$ *are now samples drawn from* p *instead of* q



In our example, using $b = 0.285$, i.e. $\widehat{P_F} = 1.5 \times 10^{-4}$. How N affects the variability of the estimator $\widehat{P_F}$?



N	$C.o.V(\widehat{P_F})$ *MCS*	$C.o.V(\widehat{P_F})$ *IS*
10	2581%	65%
100	816%	19%
1,000	258%	7% ✓
10,000	82%	2% ✓

With 1,000 samples achieves the same accuracy than 1,000,000 samples using CMCS

Co-funded by the Erasmus+ Programme of the European Union

+

- *Reduce dramatically the number of samples required by using CMCS*

−

- *An incorrect selection of the importance sampling density function can lead to erroneous estimates of the reliability*

AALBORG UNIVERSITET

Co-funded by the Erasmus+ Programme of the European Union

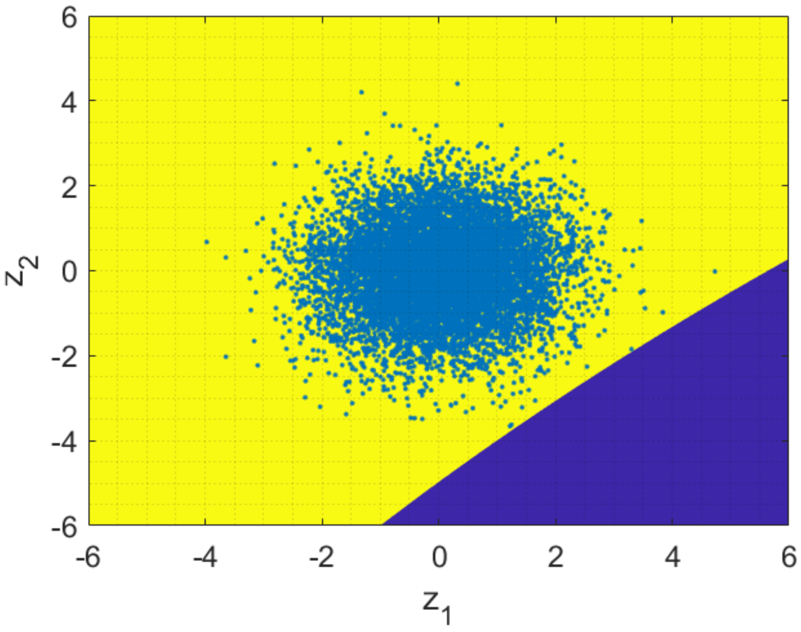
The idea is to use asymptotic properties with a simple regression technique

- *CMCS is used to estimate the failure probability in each support point*
- *The standard deviation for each support point is increased by a factor* $\frac{1}{f}$
- *Due to the larger standard deviation the failure probability is increased reducing the computational cost*
- *The functional form is given by*

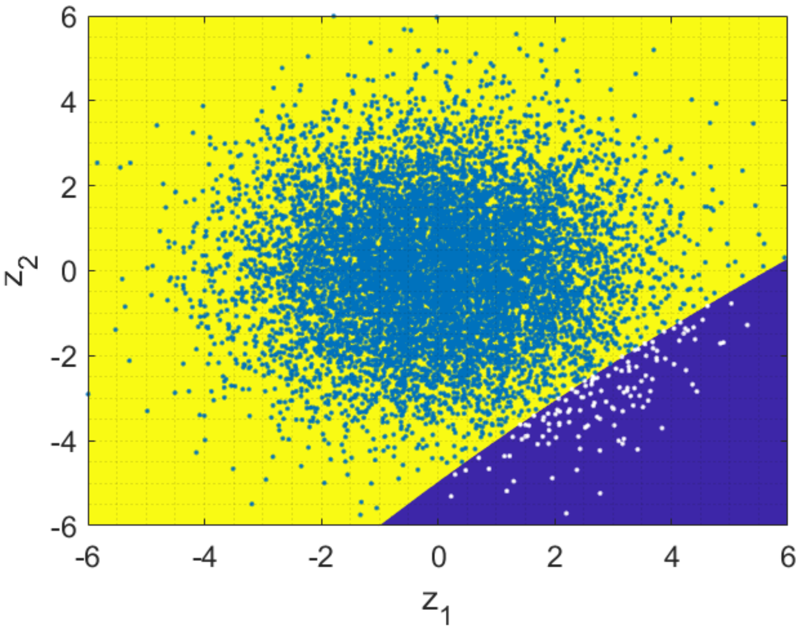
$$\beta(f) = A \cdot f + \frac{B}{f^C}$$

Co-funded by the Erasmus+ Programme of the European Union

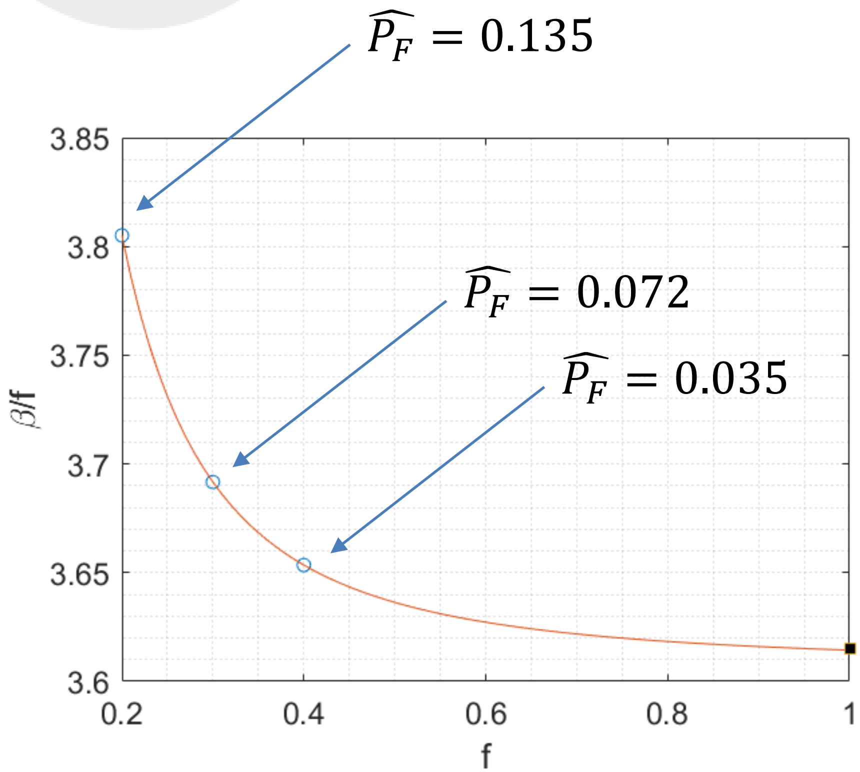
$f = 1$



$f = 0.6$



Co-funded by the
Erasmus+ Programme
of the European Union



The failure probabilities are much larger than the original implying a lower computational cost

Performing a regression, the following parameters are estimated

$$A = 3.6078$$
$$B = 0.0066$$
$$C = 2.1087$$

Then the functional form

$$\beta(f) = A \cdot f + \frac{B}{f^C}$$

Is evaluated for f=1

$$\beta(1) = 3.61$$

And then the estimator is given by

$$\widehat{P_F} = \Phi(-\beta(1)) = 1.5 \times 10^{-4}$$

Co-funded by the
Erasmus+ Programme
of the European Union

K-FORCE

Co-funded by the
Erasmus+ Programme
of the European Union

Thank you
for your attention

jgsa@civil.aau.dk

Knowledge FOr Resilient soCiEty