Uncertainty Quantification

and Reliability Analysis

Henrique Kroetz

December 15, 2025

The Deterministic Paradigm

In classical modeling, we assume inputs are known with certainty.

$$\mathcal{M}(\mathbf{x}) = f(x_1, x_2, \ldots, x_n)$$

- **Inputs (x):** Exact point values (e.g., Modulus $E = 210$ GPa).
- **Model ($\mathcal{M}$):** Physical laws (FEA, ODEs, PDEs).
- **Output (y):** A single scalar or field.

To account for unknown variations, we use arbitrary **Safety Factors** (e.g., $FS = 1.5$). This creates a binary decision (Safe/Fail) without quantifying the *margin* or likelihood of failure.

The Probability Triple: 1. The Sample Space (Ω)

To formalize uncertainty, we start with the **Sample Space**.

- The sample space Ω is the set of all possible samples or elementary events ω.
- It is defined as:

$$\Omega = \{\omega \mid \omega \in \Omega\}$$

- **In UQ terms:** This represents the universe of all possible physical realities or outcomes of an experiment (e.g., every possible microscopic configuration of a material).

The Probability Triple: 2. The Event Space ($\mathcal{F}$)

The second component is the **Event Space** (or σ-algebra).

- The σ-algebra $\mathcal{F}$ is the set of all of the considered events A.
- These events are subsets of the sample space Ω.
- It is defined as:

$$\mathcal{F} = \{A \mid A \subseteq \Omega, A \in \mathcal{F}\}$$

- **In UQ terms:** These are the macroscopic outcomes we can actually observe or measure (e.g., "the beam fails," "the temperature is above 100°C"). We cannot measure every individual ω, only sets of them.

The Probability Triple: 3. The Probability Measure (P)

The third component is the **Probability Measure**.

- The probability measure P is a function that assigns a probability $P(A)$ to every event in the event space $\mathcal{F}$.
- It maps events to a number between 0 and 1:

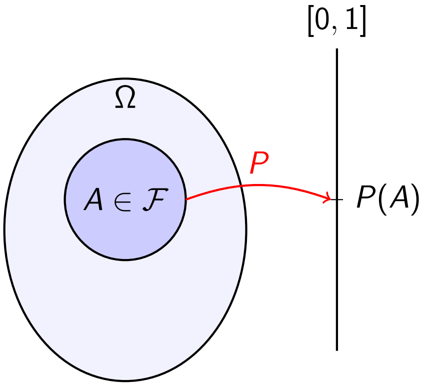
$$P : \mathcal{F} \to [0, 1]$$

- **In UQ terms:** This is how we quantify the likelihood of an observed event (e.g., the probability that the beam fails is 0.01)

Summary: The Probability Triple (or Probability Space) $(\Omega, \mathcal{F}, P)$

We define our probabilistic foundation by combining these three elements:

1. Ω **(Sample Space):**
   The set of all elementary outcomes.
2. $\mathcal{F}$ **(Event Space):**
   The collection of measurable events $A \subseteq \Omega$.
3. P **(Measure):**
   The rule assigning a number between 0 and 1 to each event in $\mathcal{F}$.



The triple $(\Omega, \mathcal{F}, P)$ completely describes the uncertainty of the system.

Random Variables: An Intuitive Introduction

What is the magnitude of the distributed load on this beam?

What is the maximum bending moment resisted by this beam?

Formal Definition

"A random variable is not random and not a variable". It is a function.

Definition

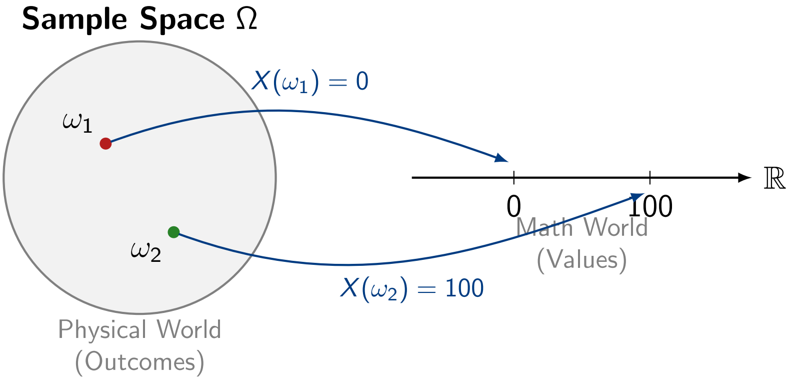
Let Ω be the sample space of a random experiment. A Random Variable X is a function that maps outcomes $\omega \in \Omega$ to the real number line $\mathbb{R}$.

$$X : \Omega \to \mathbb{R}$$

Notation:

- **Capital** X**:** The random function.
- **Lowercase** x**:** A specific realization.

Visualizing the Mapping



Discrete Random Variables: The PMF

A Random Variable (RV) is discrete when the range (set of possible values) is countable (finite or countably infinite). For a discrete variable, we define the **Probability Mass Function (PMF)**.

$$p_X(x) = \mathbb{P}(X = x)$$

Discrete Random Variables: The PMF

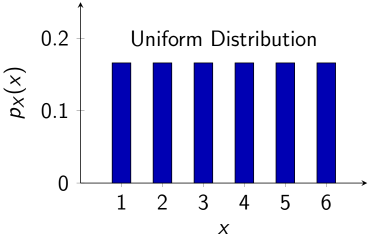
A Random Variable (RV) is discrete when the range (set of possible values) is countable (finite or countably infinite). For a discrete variable, we define the **Probability Mass Function (PMF)**.

$$p_X(x) = \mathbb{P}(X = x)$$

Example: A Fair Die

- Sample space: $\Omega = \{1, 2, 3, 4, 5, 6\}$
- Probability: $1/6 \approx 0.167$ for all x.
- The sum of all bars must equal 1.

$$\sum_i p_X(x_i) = 1$$

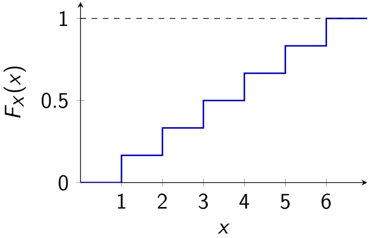


The Discrete CDF (Step Function)

The **Cumulative Distribution Function (CDF)** accumulates probability as we move from left to right.

$$F_X(x) = \mathbb{P}(X \leq x) = \sum_{x_i \leq x} p_X(x_i)$$

- It starts at 0 and ends at 1.
- For discrete variables, it looks like a staircase.
- The size of the "jump" at x equals the probability mass $p_X(x)$.



Example

Given the probability mass function (PMF) values below, calculate the Cumulative Distribution Function (CDF) and sketch its graph.

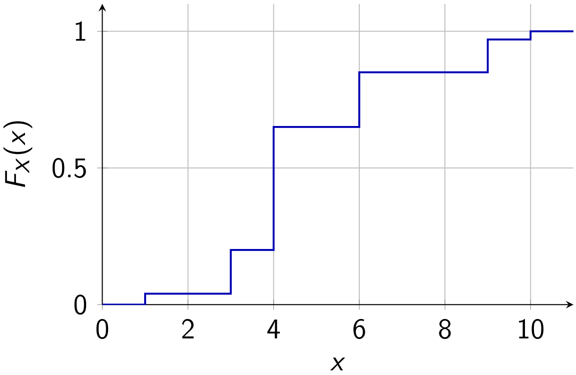
x_k	$p_X(x_k)$
1	0.04
3	0.16
4	0.45
6	0.20
9	0.12
10	0.03

Recall:

$$F_X(x) = \sum_{x_i \leq x} p_X(x_i)$$

Example

x_k	p_X	$F_X(x_k)$ (Sum)
1	0.04	$\mathbf{0.04}$
3	0.16	$0.04 + 0.16 = \mathbf{0.20}$
4	0.45	$0.20 + 0.45 = \mathbf{0.65}$
6	0.20	$0.65 + 0.20 = \mathbf{0.85}$
9	0.12	$0.85 + 0.12 = \mathbf{0.97}$
10	0.03	$0.97 + 0.03 = \mathbf{1.00}$



Discrete vs. Continuous

The type of Random Variable X is defined by the **set of values** it can take:

Discrete Random Variables

- Takes values from a **countable set** (distinct steps, e.g., integers).
- **Example:** Rolling a die, number of cars on a bridge.

Continuous Random Variables

- Takes values from a **continuum** (any value within an interval).
- **Example:** Material strength, wind speed, time.

In UQ, most physical parameters are modeled as continuous.

Defining Continuous Random Variables

1. The PDF (f_X): Density

$$f_X(x) = \frac{dF_X(x)}{dx}$$

- **Not** a probability ($f_X(x)$ can be > 1).
- $P(X = x) = 0$.
- Probability is the **Area**:

$$P(a < X < b) = \int_a^b f_X(x)dx$$

2. The CDF (F_X): Accumulation

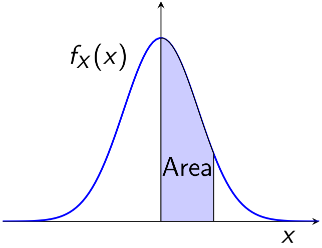
$$F_X(x) = \int_{-\infty}^{x} f_X(u)du$$

- Always a probability ($0 \leq F_X \leq 1$).
- Monotonically increasing.
- Returns the probability of being "up to" x.

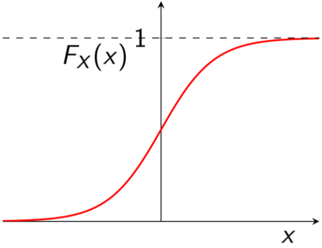
Visualizing Continuous Variables

The PDF (Density) determines the shape; the CDF (Accumulation) is the running total.

PDF (Slope of CDF)



CDF (Area of PDF)



Properties of the CDF

Regardless of whether X is discrete or continuous, the CDF $F_X(x)$ always satisfies:

1. **Bounds:** $0 \leq F_X(x) \leq 1$
2. **Monotonicity:** If $x_1 < x_2$, then $F_X(x_1) \leq F_X(x_2)$.
3. **Limits:**

$$\lim_{x \to -\infty} F_X(x) = 0 \quad \text{and} \quad \lim_{x \to +\infty} F_X(x) = 1$$

4. **Calculating Interval Probability:**

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

The CDF provides a unified way to describe any random variable.

First Moment: Expectation (Mean)

The **First Raw Moment** provides the location (center of gravity) of the distribution.

$$\mu_X = E[X^1] = \int_{-\infty}^{\infty} x f_X(x) dx$$

Physical Meaning: Center of mass.

Linearity Property: The expectation operator is linear. For constants a, b and RVs X, Y:

$$E[aX + bY + c] = aE[X] + bE[Y] + c$$

Note: This holds regardless of whether X and Y are independent.

Distinction: Raw vs. Central Moments

In Uncertainty Quantification, we distinguish between moments calculated about the origin (Raw) and moments calculated about the mean (Central).

- **Raw Moments (m_k):** Measures of magnitude/geometry.

$$m_k = E[X^k]$$

- **Central Moments (μ_k):** Measures of shape/uncertainty.

$$\mu_k = E[(X - \mu_X)^k]$$

Key Relationship:

- The **1st Raw Moment** is the Mean (μ_X).
- The **1st Central Moment** is always 0.
- The **2nd Central Moment** is the Variance.

Second Central Moment: Variance

The **Variance** is the **Second Central Moment**. It measures the spread (uncertainty) around the mean.

$$\mathrm{Var}[X] = \sigma_X^2 = E[(X - \mu_X)^2]$$

Relation to Raw Moments (Steiner's Translation): A computational shortcut relating the 2nd central moment to the 2nd raw moment:

$$\mathrm{Var}[X] = \underbrace{E[X^2]}_{\text{2nd Raw Moment}} - (\underbrace{E[X]}_{\text{1st Raw Moment}})^2$$

Properties:

- $\mathrm{Var}[X] \geq 0$
- $\mathrm{Var}[aX + b] = a^2\mathrm{Var}[X]$ (Shift b does not affect spread).
- $\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$ **only if** uncorrelated/independent.

Coefficient of Variation (CoV)

Standard Deviation depends on units. We need a dimensionless measure of uncertainty.

$$\delta_X = \text{CoV} = \frac{\sigma_X}{\mu_X}$$

Typical Engineering Values:

- **Geometric dimensions (Steel):** $\delta \approx 1 - 3\%$ (High precision).
- **Yield Strength (Steel):** $\delta \approx 5 - 10\%$.
- **Concrete Strength:** $\delta \approx 15 - 20\%$.
- **Soil Properties:** $\delta \approx 30 - 50\%$ (High uncertainty).

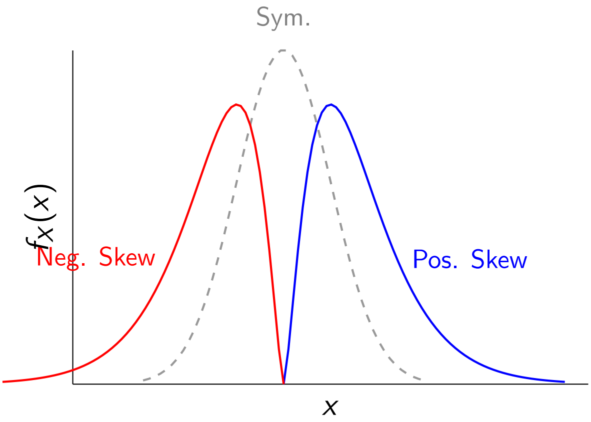
3rd Central Moment: Skewness (γ)

Measure of Asymmetry

$$\gamma = \frac{E[(X-\mu)^3]}{\sigma^3}$$

Interpretation:

- $\gamma = 0$**:** Symmetric (e.g., Normal).
- $\gamma > 0$**:** Right Skewed. Tail extends right.
- $\gamma < 0$**:** Left Skewed. Tail extends left.



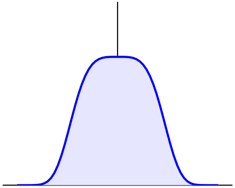
4th Central Moment: Kurtosis (κ)

Measure of "Tailedness"

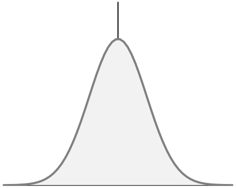
$$\kappa = \frac{E[(X-\mu)^4]}{\sigma^4}$$

- High kurtosis implies higher risk of extreme failure events than a Normal model predicts.

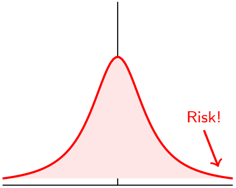
Platykurtic ($\kappa < 3$)
"Thin Tails / Flat Peak"



Mesokurtic ($\kappa = 3$)
"The Gaussian Baseline"



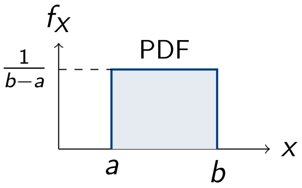
Leptokurtic ($\kappa > 3$)
"Fat Tails / Sharp Peak"



Uniform Distribution $\mathcal{U}(a, b)$

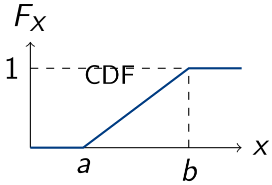
- **PDF**:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$



- **CDF**:

$$F_X(x) = \frac{x-a}{b-a} \quad \text{for } x \in [a, b]$$



Moments

- Mean: $\mu = \frac{a+b}{2}$
- Variance: $\sigma^2 = \frac{(b-a)^2}{12}$

1. Uniform Distribution

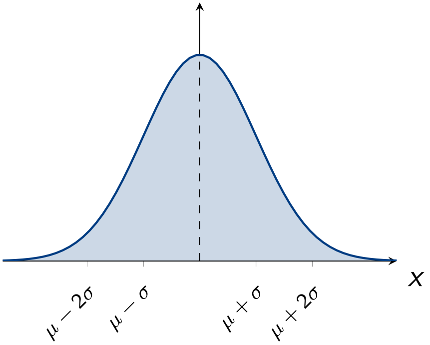
When we lack data and only have physical bounds $[a, b]$, the Uniform distribution is the most honest assumption.

- **Principle of Maximum Entropy (MaxEnt):**
  - Among all continuous distributions supported on $[a, b]$, the Uniform distribution maximizes the Entropy.
  - It assumes the "least amount of information" possible. Any other shape (e.g., triangular, Gaussian) implies knowledge of a central tendency we do not possess.

Gaussian (Normal) Distribution

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

- **Central Limit Theorem:** Sum of many independent random variables converges to a normal distribution, provided no single variable dominates the sum.
- **Examples:** Measurement noise, manufacturing tolerances, and financial stock returns.



peak density is $\approx \frac{0.4}{\sigma}$.

The Standard Normal

Any normal variable $X \sim \mathcal{N}(\mu, \sigma^2)$ can be transformed into the Standard Normal U (sometimes called Z):

$$U = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

Why?

- Simplifies integrals.
- Used in FORM/SORM reliability methods to map all variables to a common "Standard Normal Space".

Linear Combination of Gaussians

If X_1 and X_2 are Normal (even if correlated), then any linear combination $Y = aX_1 + bX_2$ is **also Normal**.

$$\mu_Y = a\mu_1 + b\mu_2$$

$$\sigma_Y^2 = a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma_2$$

- linear models with Gaussian inputs are trivial to solve.

3. Lognormal Distribution $\mathcal{LN}(\mu, \sigma^2)$

- X is Lognormal if its natural logarithm is Normally distributed:

$$\ln(X) \sim \mathcal{N}(\mu, \sigma^2)$$

- Support: $x \in (0, \infty)$ (Strictly positive).

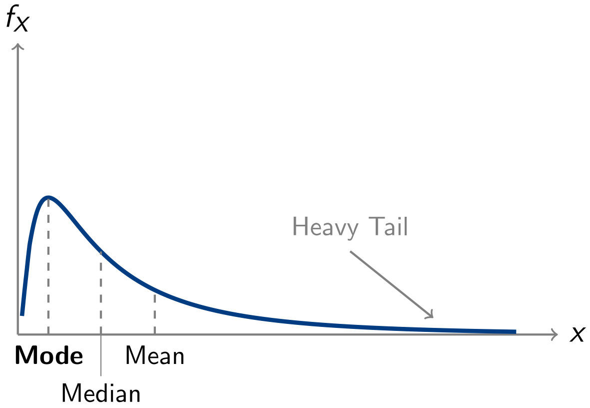
Probability Density Function

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right)$$

Real Moments (Note: $E[X] \neq \mu$!)

- Mean: $E[X] = \exp\left(\mu + \frac{\sigma^2}{2}\right)$
- Variance: $\mathrm{Var}(X) = \left(e^{\sigma^2} - 1\right) e^{2\mu+\sigma^2}$

Lognormal Distribution



- Note how the **Mean** is pulled far to the right of the **Median**.

3. Lognormal: The "Multiplicative CLT"

The Multiplicative Central Limit Theorem

If a random variable Y is the product of many independent positive random factors:

$$Y = X_1 \cdot X_2 \cdot \dots \cdot X_n$$

Then $\ln(Y) = \sum \ln(X_i)$. By the standard CLT, the sum tends to Normal $\implies$ the product tends to **Lognormal**.

Common UQ Applications:

- **Material Properties:** Fatigue life, permeability, corrosion depth.
- **Finance:** Stock prices (compounding returns).
- **Safety:** Pollutant concentrations.

Summary of Distributions

#	Distribuição	$f_X(x)$	p_1	p_2	p_3	p_4
0	Determinística	$\delta(x_0)$	x_0	-	-	-
1	Uniforme	$\frac{1}{b-a}$	a	b	-	-
2	Normal	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$	μ	σ	-	-
3	Log-Normal	$\frac{1}{\xi x\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln(x)-\lambda}{\xi}\right)^2\right]$	λ	ξ	-	-
4	Exponencial	$\upsilon \exp[-\upsilon(x-\varepsilon)]$	υ	-	ε	-
5	Rayleigh	$\frac{(x-\varepsilon)}{\eta^2} \exp\left[-\frac{1}{2}\left(\frac{x-\varepsilon}{\eta}\right)^2\right]$	η	-	ε	-
6	Logística	$\frac{e^{\frac{\pi}{\sqrt{3}}\frac{(x-\mu)}{\sigma}}}{\left(1+e^{\frac{\pi}{\sqrt{3}}\frac{(x-\mu)}{\sigma}}\right)^2}$	μ	σ	-	-
7	Gumbel mínimos	$\beta \exp[\beta(x-u_1) - e^{\beta(x-u_1)}$	u_1	β	-	-
8	Gumbel máximos	$\beta \exp[-\beta(x-u_n) - e^{-\beta(x-u_n)}$	u_n	β	-	-
9	Frechet mínimos	$\frac{\beta}{u_1}\left(\frac{x}{u_1}\right)^{\beta+1} \exp\left[-\left(\frac{x}{u_1}\right)^{\beta}\right]$	u_1	β	-	-
10	Frechet máximos	$\frac{\beta}{u_n}\left(\frac{u_n}{x}\right)^{\beta+1} \exp\left[-\left(\frac{u_n}{x}\right)^{\beta}\right]$	u_n	β	-	-
11	Weibull mínimos	$\frac{\beta}{u_1-\varepsilon}\left(\frac{x-\varepsilon}{u_1-\varepsilon}\right)^{\beta-1} \exp\left[-\left(\frac{x-\varepsilon}{u_1-\varepsilon}\right)^{\beta}\right]$	u_1	β	ε	-
12	Weibull máximos	$\frac{\beta}{\varepsilon-u_n}\left(\frac{\varepsilon-x}{\varepsilon-u_n}\right)^{\beta-1} \exp\left[-\left(\frac{\varepsilon-x}{\varepsilon-u_n}\right)^{\beta}\right]$	u_n	β	-	ε

Multivariate Distributions

In UQ, we rarely have a single variable. We have a random vector:

$$\mathbf{X} = [X_1, X_2, \dots, X_n]^T$$

Joint CDF:

$$F_{\mathbf{X}}(x_1, \dots, x_n) = \mathbb{P}(X_1 \leq x_1 \cap \dots \cap X_n \leq x_n)$$

Joint PDF:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\partial^n F_{\mathbf{X}}}{\partial x_1 \dots \partial x_n}$$

Volume under the surface $f_{\mathbf{X}}$ must equal 1.

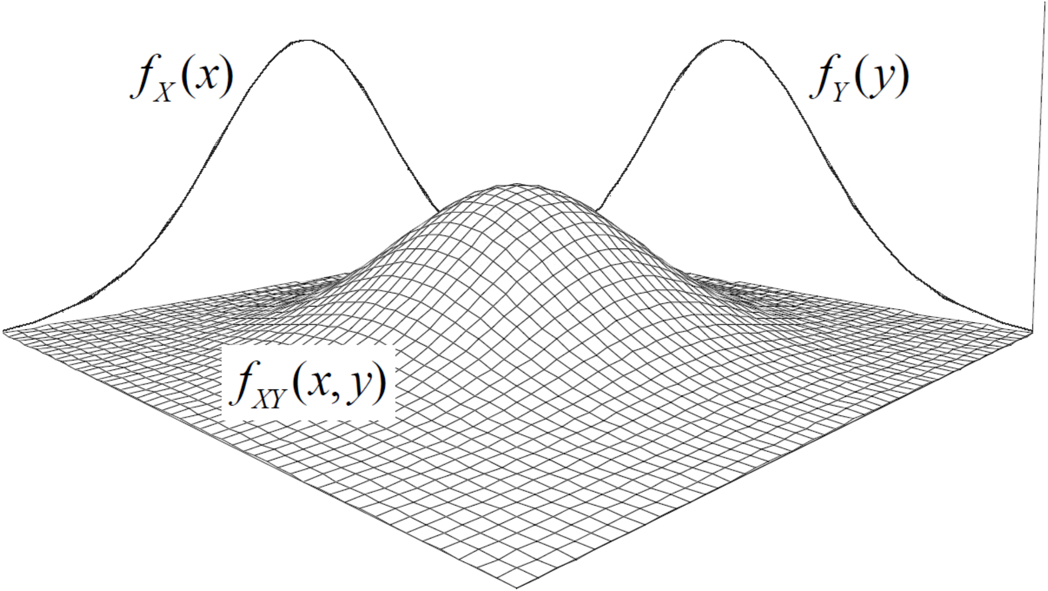
Marginalization

How do we recover the distribution of a single variable X from the joint density $f_{X,Y}$?
We "integrate out" the other variables.

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dy$$

Analogy: It's like projecting the 3D probability mountain onto one of the walls.

Visualization



Independence

Two random variables X_1 and X_2 are **statistically independent** if knowledge of one gives no information about the other.

Mathematical Condition:

$$f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2)$$

This implies that the joint density is just the product of the marginal densities. This drastically simplifies UQ problems.

Covariance

A measure of the joint variability of two random variables, indicating the direction (positive or negative) of their **linear** relationship.

$$\mathrm{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$
$$\mathrm{Cov}(X, Y) = E[XY] - E[X]E[Y]$$

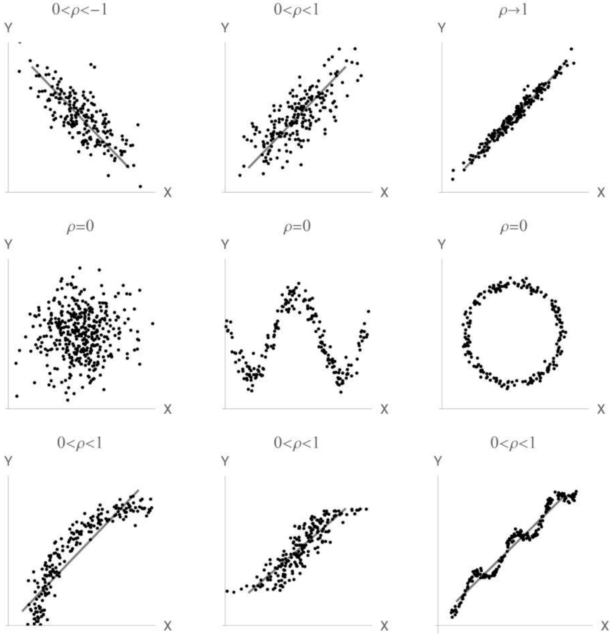
- Cov > 0: Variables tend to increase together.
  (e.g., Span length and beam weight)
- Cov < 0: One variable tends to decrease as the other increases.
  (e.g., Beam stiffness and deflection)
- Cov $= 0$: Uncorrelated.
  (Necessary but not sufficient for independence)

Correlation Coefficient (ρ)

Covariance is scale-dependent and difficult to interpret. The correlation coefficient normalizes this to provide a **dimensionless** measure of linear dependence.

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

- **Range:** $-1 \leq \rho \leq 1$
- $\rho = +1$: Perfect positive linear relationship.
- $\rho = -1$: Perfect negative linear relationship.
- $\rho = 0$: No **linear** relationship.



$0<\rho<-1$
$0<\rho<1$
$\rho\to1$
$\rho=0$
$\rho=0$
$\rho=0$
$0<\rho<1$
$0<\rho<1$
$0<\rho<1$
Y
X

The Covariance Matrix

For a random vector $\mathbf{X} = [X_1, \ldots, X_n]^T$, the covariance matrix Σ (or $\mathbf{C_X}$) is defined as the expectation of the outer product:

$$\Sigma = E[(\mathbf{X} - \mu_{\mathbf{X}})(\mathbf{X} - \mu_{\mathbf{X}})^T]$$

In component form:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \ldots & \rho_{1n}\sigma_1\sigma_n \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1n}\sigma_1\sigma_n & \ldots & \ldots & \sigma_n^2 \end{bmatrix}$$

Key Properties:

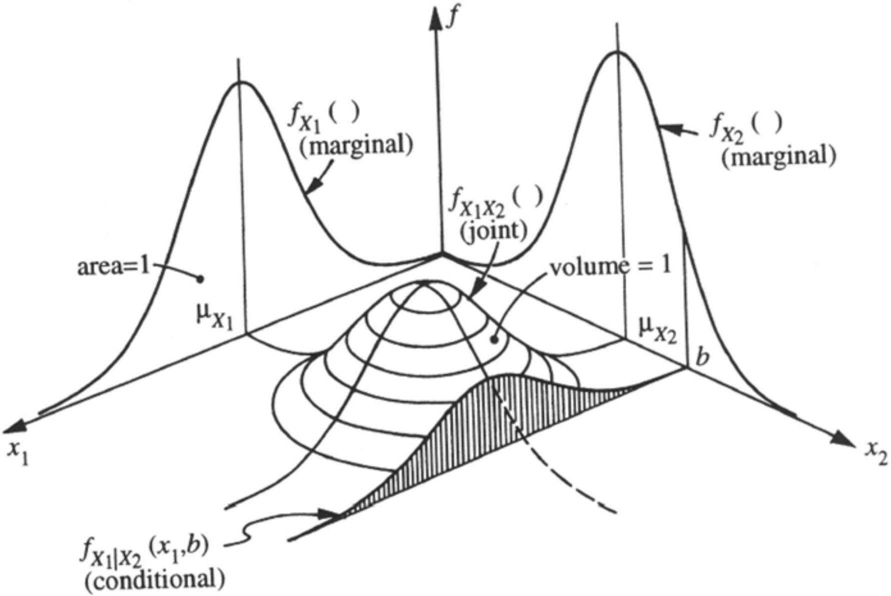
- **Symmetric:** $\Sigma = \Sigma^T$ (since $\rho_{ij} = \rho_{ji}$).
- **Diagonal:** Elements are variances (σ_i^2).
- **Positive Semi-Definite:** $\mathbf{y}^T \Sigma \mathbf{y} \geq 0$ for any vector $\mathbf{y}$.

The Multivariate Normal Distribution (MVN)

The fundamental model for coupled uncertainties in engineering. If $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, its PDF is:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n|\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}\underbrace{(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}_{\text{Mahalanobis Distance}^2}\right)$$

- $\boldsymbol{\mu}$ defines the center; $\boldsymbol{\Sigma}$ defines the shape (contours are ellipsoids).
- All marginals and conditionals (slices) are also Gaussian.
- For Gaussians, uncorrelated ($\rho = 0$) implies Independence.



f
$f_{X_1}(\ )$
(marginal)
$f_{X_2}(\ )$
(marginal)
$f_{X_1X_2}(\ )$
(joint)
area=1
volume = 1
μ_{X_1}
μ_{X_2}
b
x_1
x_2
$f_{X_1|X_2}(x_1,b)$
(conditional)