# Statistics and Probability Theory

# in

# Civil, Surveying and Environmental Engineering

**Prof. Dr. Michael Havbro Faber**

**Swiss Federal Institute of Technology**

**ETH Zurich, Switzerland**

# Contents of Todays Lecture

- Overview of Uncertainty Modeling

- Random Variables
  - properties of the expectation operator
  - random vectors and joint moments
  - conditional distributions and conditional moments
  - the probability distribution for the sum of two random variables
  - the probability distribution for functions of random variables

# Overview of Uncertainty Modeling

- Random variables and their characteristics

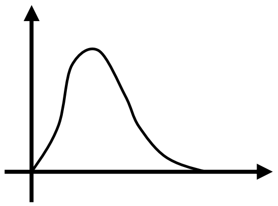
**Real world**

**Uncertain phenomenon**

↓

**Data/observations**

**Model**



**Random variables**

↓

| x,y | p(x,y) |
|---|---|
| 1,10 | 0.033 |
| 1,20 | 0.067 |
| 1,30 | 0.033 |
| 1,40 | 0.033 |
| 2,10 | 0.067 |
| 2,20 | 0.100 |
| 2,30 | 0.067 |
| 2,40 | 0.033 |
| 3,10 | 0.067 |
| 3,20 | 0.133 |
| 3,30 | 0.100 |
| 3,40 | 0.067 |
| 4,10 | 0.033 |
| 4,20 | 0.067 |
| 4,30 | 0.067 |
| 4,40 | 0.033 |

## Random Variables

- Properties of the expectation operator

  The expectation operator facilitates that we can assess the expected value and the variance of a random variable

  By understanding how the expectation operator works we will be able to assess the expected value and the variance of functions of random variables

  This is useful if we want to analyze engineering models involving one or more random variables in regard to their expected values and their variances

  E.g.: Duration of a construction process as a function of the duration of its individual processes

## Random Variables

- Properties of the expectation operator

The expectation operator possesses the following properties:

$$E[c] = c$$

$$E[cX] = cE[X]$$

$$E[a + bX] = a + bE[X]$$

$$E[g_1(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)]$$

## Random Variables

- Properties of the expectation operator

The variance can thus be written as:

$$
\begin{aligned}
Var[X] &= E\left[(X-\mu_X)^2\right] \\
&= E\left[X^2+\mu_X^2-2\mu_X X\right] \\
&= \mu_X^2+E\left[X^2\right]-2\mu_X E[X] \\
&= \mu_X^2+E\left[X^2\right]-2\mu_X^2 = E\left[X^2\right]-\mu_X^2
\end{aligned}
$$

# Random Variables

- Properties of the expectation operator

$$E[c] = c$$
$$E[cX] = cE[X]$$
$$E[a + bX] = a + bE[X]$$
$$E[g_1(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)]$$

Furthermore there is

$$Var[c] = 0$$
$$Var[cX] = c^2 Var[X]$$
$$Var[a + bX] = b^2 Var[X]$$

# Random Variables

- Properties of the expectation operator

From the result

$$Var[X] = E\left[(X-\mu_X)^2\right] = E\left[X^2 + \mu_X^2 - 2\mu_X X\right] = E\left[X^2\right] - \mu_X^2$$

it is seen that there in general is $E[g(X)] \neq g(E[X])$

$E[g(X)] \geq g(E[X])$ for convex functions - Jensen's inequality !

**Equality only for linear functions**

## Random Variables

- **Random vectors and joint moments**

**Often we are dealing with models involving not only one random variable but several random variables**

**These random variables can be collected in a vector**

**In general the components of the vector are dependent**

**E.g. Rainfall and water level**

**It is thus necessary that we establish probabilistic models which include this dependency – we can do this through the joint cumulative distributions and the joint moments.**

# Random Variables

- **Random vectors and joint moments**

**Now we consider not just one continuous random variable but a vector of continuous random variables**

$$\mathbf{X} = \left( X_1, X_2, \ldots, X_n \right)^T$$

**The *joint cumulative distribution function* is given by**

$$F_{\mathbf{X}}\left(\mathbf{x}\right) = P\left( X_1 \le x_1 \cap X_2 \le x_2 \cap \ldots \cap X_n \le x_n \right)$$

**and the *joint probability density function* is given by**

$$f_{\mathbf{X}}\left(\mathbf{x}\right) = \frac{\partial^n}{\partial z_1 \partial z_2 \ldots \partial z_n} F_{\mathbf{X}}\left(\mathbf{x}\right)$$

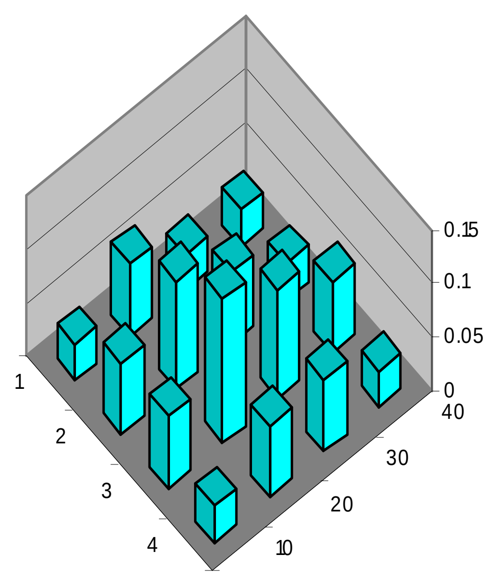
# Random Variables

- **Random vectors and joint moments**

**Consider the two dimensional discrete probability density function:**

| x,y | p(x,y) |
|---|---|
| 1,10 | 0.033 |
| 1,20 | 0.067 |
| 1,30 | 0.033 |
| 1,40 | 0.033 |
| 2,10 | 0.067 |
| 2,20 | 0.100 |
| 2,30 | 0.067 |
| 2,40 | 0.033 |
| 3,10 | 0.067 |
| 3,20 | 0.133 |
| 3,30 | 0.100 |
| 3,40 | 0.067 |
| 4,10 | 0.033 |
| 4,20 | 0.067 |
| 4,30 | 0.067 |
| 4,40 | 0.033 |

$\sum = 1$



## Random Variables

- **Random vectors and joint moments**

  **The *marginal probability density function* of a random variable $X_i$ is defined by**

$$f_{X_i}\left(x\right)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(n-1\ \text{ fold})\, f_{\mathbf{X}}\left(\mathbf{x}\right)dx_1..dx_{i-1}dx_{i+1}..dx_n$$

# Random Variables

- Random vectors and joint moments

Consider the two dimensional discrete probability density function:

| x,y | p(x,y) |
|---|---|
| 1,10 | 0.033 |
| 1,20 | 0.067 |
| 1,30 | 0.033 |
| 1,40 | 0.033 |
| 2,10 | 0.067 |
| 2,20 | 0.100 |
| 2,30 | 0.067 |
| 2,40 | 0.033 |
| 3,10 | 0.067 |
| 3,20 | 0.133 |
| 3,30 | 0.100 |
| 3,40 | 0.067 |
| 4,10 | 0.033 |
| 4,20 | 0.067 |
| 4,30 | 0.067 |
| 4,40 | 0.033 |

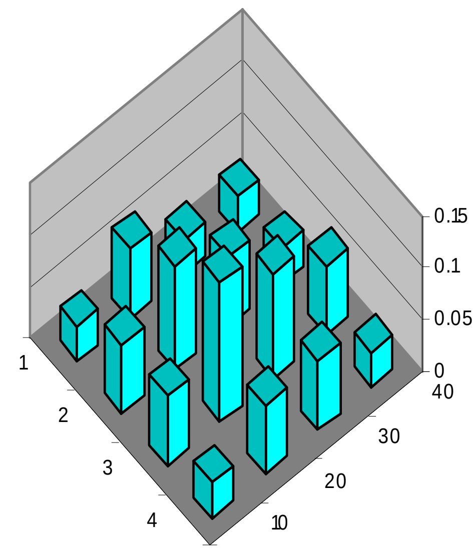
$\Sigma = 0.17$

$\Sigma = 0.27$

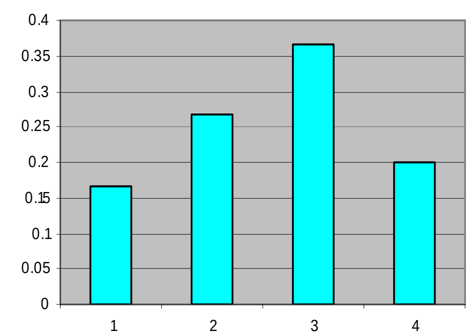
$\Sigma = 0.37$

$\Sigma = 0.2$

$\Sigma = 1$




**Discrete joint density**




**Marginal density for x**

## Random Variables

- Random vectors and joint moments

The *covariance* between the i'th and the j'th component of the random vector of continuous random variables is defined as the *joint central moment* i.e. by

$$C_{X_iX_j} = E\left[(X_i-\mu_{X_i})(X_j-\mu_{X_j})\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(x_i-\mu_{X_i}\right)\left(x_j-\mu_{X_j}\right) f_{X_iX_j}\left(x_i,x_j\right)dx_i dx_j$$

$$C_{X_iX_i} = Var\left[X_i\right]$$

From where we see that for i = j we get the variance for $X_i$

**Correlation coefficient** $\rho_{X_iX_j} = \dfrac{C_{X_iX_j}}{\sigma_{X_i}\sigma_{X_j}}$ $\rho_{X_iX_i} = 1$

## Random Variables

- Random vectors and joint moments

  The expected value and the variance of a linear function

$$Y = a_0 + \sum_{i=1}^{n} a_i X_i$$

are given by

$$E[Y] = a_0 + \sum_{i=1}^{n} a_i E[X_i]$$

$$Var[Y] = \sum_{i=1}^{n} a_i^2 Var[X_i] + \sum_{\substack{i,j=1 \\ i \neq j}}^{n} a_i a_j C_{X_i X_j}$$

## Random Variables

- Conditional distributions and conditional moments

Some times it is useful to be able to assess the probability of an event given that we know something about one of the random variables which are used to define the event

E.g. assume we want to calculate the probability that a project will be delayed under the condition that one of the processes will exceed its planned duration by 50%.

## Random Variables

- **Conditional distributions and conditional moments**

The *conditional probability density function* for the random variable $X_1$ given the outcome of the random variable $X_2$ is given by

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

where if $X_1$ and $X_2$ are independent

$$f_{X_1|X_2}(x_1|x_2) = f_{X_1}(x_1)$$

The *conditional cumulative distribution function* is obtained by integration as

$$F_{X_1|X_2}(x_1|x_2) = \frac{\int_{-\infty}^{x_1} f_{X_1,X_2}(z,x_2)dz}{f_{X_2}(x_2)}$$

# Random Variables

- **Conditional distributions and conditional moments**

  **The *un-conditional cumulative distribution function* for the random variable *$X_1$* can be derived from the conditional comulative distribution function by use of the *total probability theorem***

$$F_{X_1}(x_1) = \int_{-\infty}^{\infty} F_{X_1|X_2}(x_1|x_2) f_{X_2}(x_2) dx_2$$

**The *conditional expected value* is defined by**

$$\mu_{X_1|X_2} = E\left[ X_1 | X_2 = x_2 \right] = \int_{-\infty}^{\infty} x_1 \, f_{X_1|X_2}\left( x|x_2 \right) dx_1$$

## Random Variables

- In many cases we are interested in assessing the probabilites of functions of random variables

  The functions are useful for describing the events we are interested in – they are our engineering models.

  A simple case is the sum of two random variables – it is useful to derive the cumulative distribution function for such a sum.

  A more general case concerns monotonic functions of random variables – we will also derive the cumulative distribution for this case.

## Random Variables

- The cumulative distribution function for the sum of two random variables

  Consider the sum $Y = X_1 + X_2$

  and assume that we have $f_{X_1,X_2}(x_1,x_2)$

  First we derive the density function for $Y = x_1 + X_2$

  assuming that $X_1$ is given i.e. $f_{X_2|X_1}(x_2|x_1) = \dfrac{f_{X_1,X_2}(x_1,x_2)}{f_{X_1}(x_1)}$

  $$f_{Y|X_1}(y|x_1) = f_{X_2|X_1}(y - x_1|x_1)$$

  and we get $f_{Y,X_1}(y,x_1) = f_{X_2|X_1}(y - x_1|x_1)\, f_{X_1}(x_1) = f_{X_2,X_1}(y - x_1, x_1)$

## Random Variables

- The cumulative distribution function for the sum of two random variables

  The marginal probability density function for *Y* is now achieved by integrating out over $X_1$, i.e.

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_2,X_1}(y - x_1, x_1)\,dx_1$$

  For the case where $X_1$ and $X_2$ are independent we get the so-called *convolution integral*

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_2}(y - x_1) f_{X_1}(x_1)\,dx_1$$

## Random Variables

- The cumulative distribution function for functions of random variables

  Consider the more general problem of deriving the cumulative distribution function for a function of a random variables i.e. $Y = g(X)$ where the probability distribution function of $X$ is given as $F_X(x)$

  If $g(x)$ is monotonically increasing and represents a one-to-one mapping, a realization of $Y$ is only smaller than $y_0$ if the realization of $X$ is smaller than $x_0$ where $x_0 = g^{-1}(y_0)$

$$F_Y(y) = P(Y \leq y) = P(X \leq g^{-1}(y))$$

The cumulative distribution function for Y is then given by

$$F_Y(y) = F_X(g^{-1}(y))$$

## Random Variables

- The cumulative distribution function for functions of random variables

starting now with $F_Y(y) = F_X(g^{-1}(y))$

we have

$$f_Y(y) = \frac{\partial F_X(g^{-1}(y))}{\partial y}$$

$$f_Y(y) = \frac{\partial}{\partial y} g^{-1}(y) f_X(g^{-1}(y)) \quad \Rightarrow \quad f_Y(y) = \frac{\partial x}{\partial y} f_X(x)$$

## Random Variables

- The cumulative distribution function for functions of random variables

  In case the function $g(x)$ is monotonically decreasing, a realization of *Y* is only smaller than $y_0$ if the realization of *X* is larger than $x_0$ , and in this case we have to change the sign i.e.

$$F_Y(y) = -F_X(g^{-1}(y))$$

  yielding $f_Y(y) = -\frac{\partial x}{\partial y} f_X(x)$

  In the general case – for monotonically increasing or decreasing functions there is thus

$$f_Y(y) = \left|\frac{\partial x}{\partial y}\right| f_X(x)$$

## Random Variables

- The cumulative distribution function for functions of random variables

  For the case where the components of a random vector $\mathbf{Y} = (Y_1, Y_2, ..Y_n)^T$ can be given as one-to-one mappings of monotonically increasing or decreasing functions $g_i, i = 1, 2, ..n$ of the components of a random vector $\mathbf{X} = (X_1, X_2, ..X_n)^T$

  in the form: $Y_i = g_i(\mathbf{X})$

  there is $f_{\mathbf{Y}}(\mathbf{y}) = |\mathbf{J}| \, f_{\mathbf{X}}(\mathbf{x})$

  with $|\mathbf{J}|$ being the absolute value of the determinant of

$$\mathbf{J} = \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & ..... & \frac{\partial x_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & ..... & \frac{\partial x_n}{\partial_n} \end{bmatrix}$$