CV 510$^{0}_{1}$
Modeling, Uncertainty, and Data for Engineers
(July – Nov 2025)

Dr. Prakash S Badal

1

Flow

- Announcement
- Module Outline

2

2

Announcement

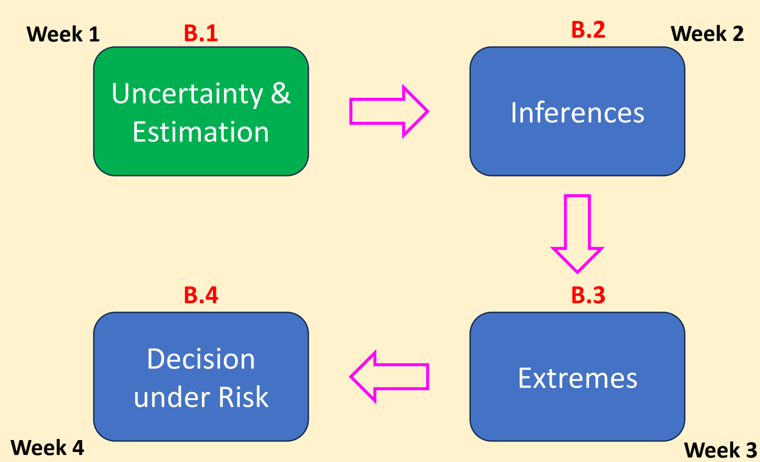
- This week's classes
  - Thursday (11th Sep) usual class hour
  - Additional class on Friday (12th Sep at 8 AM)

- Check GitHub regularly

3

3

Module Overview



A crash course in probability, probabilistic models, and probabilistic methods.

4

Check-in with teach-book

https://mude.citg.tudelft.nl/book/2024/

- Q1 Topics (Chapters 5, 6, 2, and 3)
  - Q1C5 Univariate continuous distribution
    - Q1C5.1 PDF/CDF
    - Q1C5.2 Empirical Distributions
    - Q1C5.3 Parametric Distributions
    - Q1C5.4 Fitting a Distribution
  - Q1C6 Multivariate Distributions (briefly)
  - Q1C2 Propagation of Uncertainty
  - Q1C3 Observation Theory: least-sq., Hyp. Test, Conf. Intervals
- Q2 Topics (Chapter 7 and 8)
  - Q2C7 Extreme value theory: GEV, return period, POT
  - Q2C8 Risk and decision making (CBA)
- Fundamental Concepts
  - Chapter 6, 7, 8, 9. Probability basics, rv, z- and t-tables
- Programming
  - Fundamental Concepts → Chapter 10

5

5

Module Overview

B.1

Uncertainty & Estimation

B.1.1

Probabilistic basics and RVs
- Probability laws
- Random variables
- Discrete, continuous RVs
- PMF/PDF/CDF
- Expectation
- Variance/cov./corr.

B.1.2

Core distributions
- Gaussian
- Uniform
- Exponential
- Lognormal
- Symmetric/skewness
- *Tails*: thick/light

B.1.3

Propagation laws & least-sq.
- Uncertainty propagation
- Least-square errors
- Parameter estimation

6

B.1.1 Probability basics & random variables

7

Settlers of Catan

[Wikimedia commons](#)

[colonist.io](#) with permission

8

8

Practice question

- Work with the three people around you and think about the following experiment.
  - You throw 2 dice; add the values of each die
  - List all the possible outcomes

Q1. What is the probability that the sum of the two dice will be 6?

Q2. What is the probability that you get at least one 8 in a single round consisting of four players?

Alternatively, browse

etc.ch/FBb6

9

Probability axioms

- If S is the sample space and E is an events in a random experiment, 3 probability axioms are

$$\Pr(S) = 1$$

$$0 \leq \Pr(E) \leq 1$$

If two events E_1 and E_2 are mutually exclusive,

$$\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2)$$

10

10

Probability rules

- Union: $\Pr(E_1 \cup E_2) =$ $\Pr(E_1) + \Pr(E_2) - \Pr(E_1E_2)$
- Complement: $\Pr(\bar{E}) =$ $1 - \Pr(E)$
- Inclusion-exclusion:

$$\Pr(E_1 \cup E_2 \cup E_3) =$$

$$\Pr(E_1) + \Pr(E_2) + \Pr(E_3) - \Pr(E_1E_2) - \Pr(E_2E_3) - \Pr(E_1E_3) + \Pr(E_1E_2E_3)$$

$$= \sum_i \Pr(E_i) - \sum_{i<j} \Pr(E_iE_j) + \sum_{i<j<k} \Pr(E_iE_jE_k)$$

11

11

Probability rules: conditional prob.

- Conditional probability: $\Pr(E_1|E_2) = \frac{\Pr(E_1E_2)}{\Pr(E_2)}$

Probability of E_1 given that E_2 has occurred.

- Multiplication rule:

$$\Pr(E_1E_2) = \Pr(E_1|E_2) \cdot \Pr(E_2)$$

Note the notation.
$\Pr(E_1 \cap E_2) \equiv \Pr(E_1E_2)$

- Two events are called statistically independent (SI), if

"Conditional probability of one event given the other has occurred" is identical to "its marginal probability."

$$\Pr(E_1|E_2) = \Pr(E_1)$$

Thus, for statistically independent events,

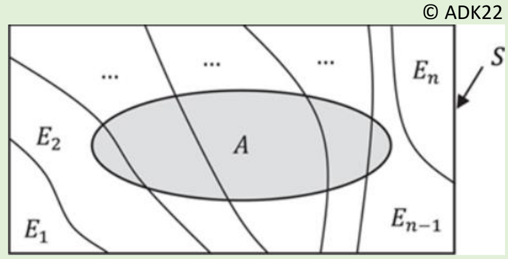
$$\Pr(E_1E_2) = \Pr(E_1) \cdot \Pr(E_2)$$

12

12

Probability rules: total probability rule

- If $E_1, E_2, \ldots, E_n$ are n MECE events,



$\Pr(A) =$

$\Pr(AE_1) + \Pr(AE_2) + \cdots + \Pr(AE_n)$

$= \sum_{i=1}^{n} \Pr(AE_i)$

$= \sum_{i=1}^{n} \Pr(A|E_i)\Pr(E_i)$

- Breaking down of calculation of event A into computing the conditional probabilities $P(A|E_i)$
- Conditionals usually easier to compute
- Clever selection of events E_i

13

13

(At home) Probability rules: total probability rule

The probability that a building will sustain damage due to an earthquake depends on the intensity of shaking. Suppose the intensity is discretized into the states of weak, moderate, and strong. Assume it is determined that the probability of damage to the building is 0.01, 0.10, and 0.60 for the three levels of shaking intensity, respectively. Furthermore, based on the history of past earthquakes in the region, it is estimated that, for a randomly occurring earthquake, the probabilities for the three levels of shaking intensity at the site of the building are 0.90, 0.08, and 0.02, respectively. We are interested in the probability that the building will sustain damage during the next earthquake in the region.

Denote earthquake intensity by W, M, and S; and damage by D.

$\Pr(W) = 0.90 \quad \Pr(M) = 0.08 \quad \Pr(S) = 0.02$

$\Pr(D|W) = 0.01 \quad \Pr(D|M) = 0.10 \quad \Pr(D|S) = 0.60$

$\Pr(D) = \Pr(D \cap W) + \Pr(D \cap M) + \Pr(D \cap S)$

$= \Pr(D|W)\Pr(W) + \Pr(D|M)P(M) + \Pr(D|S)P(S)$

$= 0.009 + 0.008 + 0.012 = 0.029$

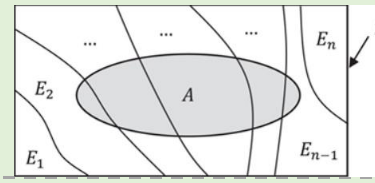
14

14

Probability rules: Bayes' rule

© ADK22

- If $E_1, E_2, \ldots, E_n$ are n MECE events,



Multiplication rule gives:

$\Pr(AE_i) = \Pr(E_i|A)\Pr(A) = \Pr(A|E_i)\Pr(E_i)$

$$\Rightarrow \Pr(E_i|A) = \frac{\Pr(A|E_i)\Pr(E_i)}{\Pr(A)}$$

$$\Rightarrow \Pr(E_i|A) = \frac{\Pr(A|E_i)\Pr(E_i)}{\sum_{j=1}^{n} \Pr(A|E_j)\Pr(E_j)}$$

15

15

(At home) Probability rules: Bayes' rule

- Assume in the previous example that the building has sustained damage. What can we learn about the intensity of the earthquake?

$\Pr(W) = 0.90$ $\quad \Pr(M) = 0.08$ $\quad \Pr(S) = 0.02$

$\Pr(D|W) = 0.01$ $\quad \Pr(D|M) = 0.10$ $\quad \Pr(D|S) = 0.60$

$Pr(D) = \Pr(D|W)\Pr(W) + \Pr(D \cap M)\Pr(M) + \Pr(D \cap S)Pr(S)$

$\quad = 0.009 + 0.008 + 0.012 = 0.029$

$\Pr(W|D) = 0.310$ $\quad \Pr(M|D) = 0.276$ $\quad \Pr(S|D) = 0.414$

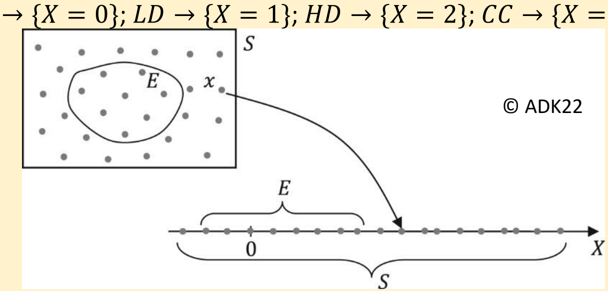
Observe the updated probabilities (~31%, 28%, 41%).
They are significantly different from the prior probabilities (90%, 8%, 2%).
Knowing that the damage has occurred, the probabilities for the intensity to have been Moderate or Strong have sharply increased.
Also, note that probabilities still add up to 1. As {W, M, S} were MECE.

16

16

Random variables

- A rv is defined by mapping of the sample space onto a line.
- Each sample point is mapped onto a point on the line.
- Sample space and events represent intervals on the line.
  - Strength of concrete cubes (natural numerical values; obvious mapping)
  - Possible states of a building after an earthquake: no damage (ND), light damage (LD), heavy damage (HD), and complete collapse (CC).
    - No natural numerical value. Define a rule
    - $ND \to \{X = 0\}; LD \to \{X = 1\}; HD \to \{X = 2\}; CC \to \{X = 3\}$



17

17

Random variable (rv)

- Notation:
  - Denote by an uppercase letter such as X
    - C = a coin toss; F_c = strength of concrete.
  - Measured values of the RV from experiments are denoted by a lowercase letter such as x
    - c = the outcome of a coin toss; f_c = experimental strength of a concrete specimen.
- Discrete rv
  - Finite (or countably infinite) range
    - Number of scratches on a surface, number of sand on a beach, proportion of defective parts among 1000 tested
- Continuous rv
  - Interval of real numbers (finite or infinite) as range
    - Concrete strength, temperature, mass

18

18

Random variable (rv): PMF, PDF

- For a discrete rv: *probability mass function* (PMF) is defined as:

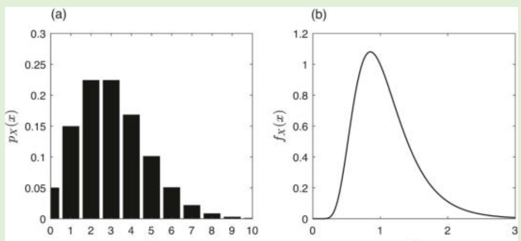
$$p_X(x) = \Pr(X = x)$$

PMF must satisfy: $0 \le p_X(x) \le 1$ and $\sum_x p_X(x) = 1$

- For a continuous rv: *probability density function* (PDF) is defined as:

$$f_X(x)\mathrm{d}x = \Pr(x < X \le x + \mathrm{d}x)$$

PDF must satisfy: $0 \le f_X(x)$ and $\int_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x = 1$



 19

19

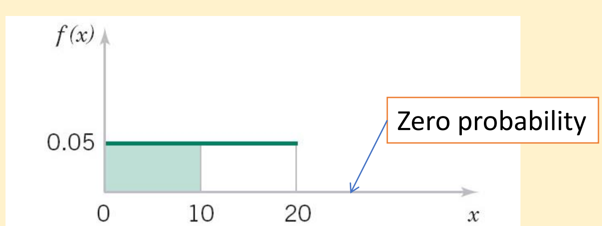
(At home) Random variable (rv): PMF, PDF

- Let X = concrete strength, in the range $[0, 20\ \mathrm{MPa}]$
- Assume a PDF given as:

$$f_X(x) = 0.05 \text{ for } 0 \le x \le 20$$

- What is the probability that the strength of a tested concrete is less than 10 MPa?

$$\Pr(X < 10) = \int_0^{10} f_X(x)\mathrm{d}x = \int_0^{10} 0.05\ \mathrm{d}x = 0.5$$



20

20

Random variable (rv): CDF

- *Cumulative distribution function* (CDF) is defined as:

$$F_X(x) = \Pr(X \le x)$$

- The CDF of a discrete RV X is

$$F_X(x) = \Pr(X \le x) = \sum_{x \le x_i} p_X(x_i)$$

- The CDF of a continuous RV X is

$$F_X(x) = \Pr(X \le x) = \int_{-\infty}^{x} f_X(x)\mathrm{d}x$$

Derivative of $F_X(x)$ for continuous rv:

$$f_X(x) = \frac{\mathrm{d}F_X(x)}{\mathrm{d}x}$$

21

21

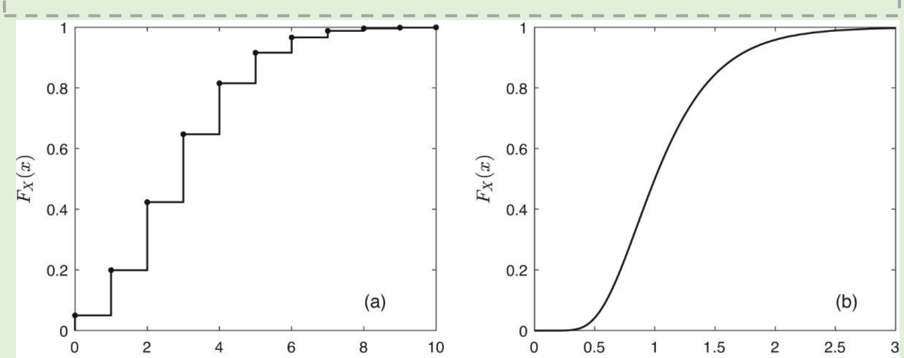
Random variable (rv): CDF

CDF must satisfy:

$F_X(-\infty) = 0$

$F_X(+\infty) = 1$

if $x \le y$, then $F_X(x) \le F_X(y)$. Non-decreasing function



 22

22

Expectation

- Let $g(X)$: a function of rv X, with PDF $f_X(x)$.
  Expectation of $g(X)$,

$$\mathrm{E}[g(X)] = \int_{-\infty}^{+\infty} g(x) f_X(x) \mathrm{d}x$$

$$\mathrm{E}[\sum_i g_i(\mathbf{X})] = \sum_i \mathrm{E}[g_i(\mathbf{X})]$$

- Expectation: weighted average of the function over all outcomes with weights as corresponding probabilities.
- Expectation: a linear operator.

- Expectation/mean of X

$$\mathrm{E}[X] = \int_{-\infty}^{+\infty} x f_X(x) \mathrm{d}x = \mu_X$$

23

23

Variance

- Variance is the expected value of $X - \mu_X$,

$$\sigma^2 = \mathrm{E}[(X-\mu_X)^2]$$
$$= \mathrm{E}\left[X_i^2 + \mu_i^2 - 2X_i\mu_i\right]$$
$$= \mathrm{E}\left[X_i^2\right] + \mathrm{E}\left[\mu_i^2\right] - \mathrm{E}[2X_i\mu_i]$$
$$= \mathrm{E}\left[X_i^2\right] + \mu_i^2 - 2\mu_i\mathrm{E}[X_i] = \mathrm{E}\left[X_i^2\right] - \mu_i^2$$

$$\Rightarrow \mathrm{Var}[X_i] = \mathrm{E}\left[X_i^2\right] - \mathrm{E}^2[X_i]$$

24

24

Covariance

- Covariance is the expected value of $(X_1 - \mu_1)(X_2 - \mu_2)$,

$\mathrm{Cov}[X_1, X_2] =$ $\mathrm{E}[(X_1 - \mu_1)(X_2 - \mu_2)]$

(at home exercise)

$\Rightarrow \mathrm{Cov}[X_1, X_2]$

$= \mathrm{E}[X_1 X_2] - \mathrm{E}[X_1]\mathrm{E}[X_2]$

25

25

Correlation

- Correlation coefficient

$$\rho_{12} = \frac{\mathrm{Cov}[X_1, X_2]}{\sigma_1 \sigma_2}.$$

$$\mathrm{Cov}[X_1, X_2] = \rho_{12}\sigma_1\sigma_2$$

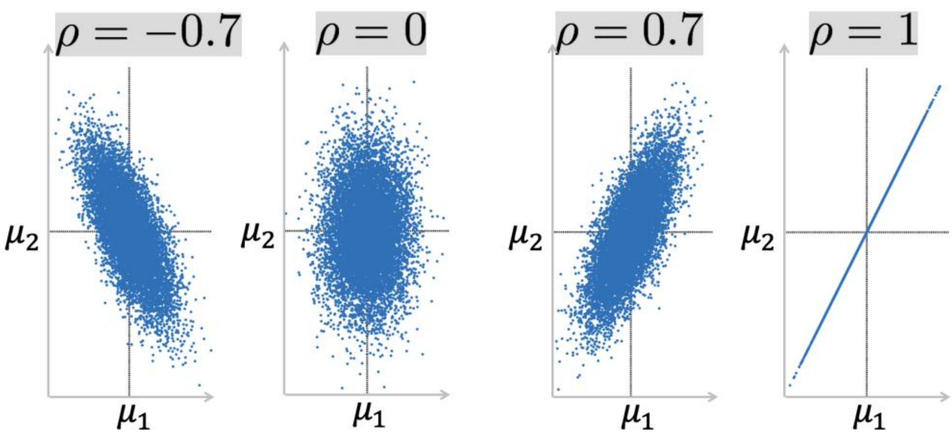
$$-1 \le \rho_{12} \le 1$$

Provides a measure of linear dependence between two rv's. When $\rho_{12} = \pm 1$, fully correlated.

26

26

Correlation



27

27

Questions, comments, or concerns?

28

28