Reliability Analysis

Module 2: Basic Reliability Math: Probability

Prof. Katrina M. Groth
Mechanical Engineering Department
Center for Risk and Reliability
University of Maryland
kgroth@umd.edu

Objectives for this module

- At the end of this module you will [be able to]:
  - Manipulate events (and data) using sets & Boolean algebra
  - Know laws of probability
  - Know and use Bayes' Theorem
  - Solve problems using common parametric probability distributions

Random variables (events) of interest

- Notation: A capital letter (e.g., X) denotes a random variable, a lower-case letter (x) denotes a value that the r.v. can take.
- X could be an r.v representing:
  - Time-to-failure of an item: $x \in [0, \infty)$
  - Number of failures occurring in some time: $x \in [0, 1, 2, \dots, n]$
  - Number of cycles until first failure occurs: $x \in [0, 1, 2, \dots, n]$
  - Number of failed components drawn from a population: $x \in [0, 1, 2, \dots, n]$
  - And more

Sets and Boolean algebra

- A **set** is a collection of items or elements, each with some specific characteristics.
- A **universal set** Ω: is a set that includes all items of interest
- A **null set or empty set ,** $\varnothing$ is a set with no items.
- A **subset,** ($\subset$, or *subset equal to* $\subseteq$) refers to a collection of items that belong to a set
- $B \subset A \subset \Omega$.

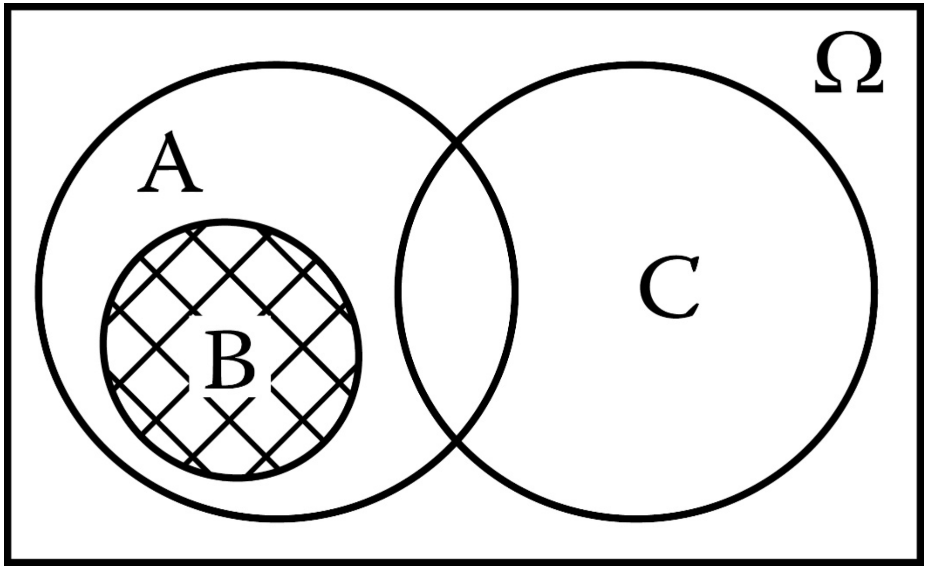
In engineering notation, Ω is often replaced by 1 and $\varnothing$ by 0.

Example sets

- Example continuous set:
  - For T, failure times for light bulbs.
    - $A = \{t \mid t > 0\}$ could show the universal set
    - $B = \{t \mid 0 < t < 100\}$ shows a subset of light bulbs that operate for $t > 0$ but fail prior to 100 hours.
- Example discrete sets:
  - States of a pump A: {Failed, working}
    - We may be interested in analyzing only those that are failed…
  - A={1,2,3,4,5,6} is the universal set of outcomes from rolling a die once.
    - B={4,5,6} is the subset of dice outcomes that are *greater than 3*.

Sets and Boolean algebra (cont.)

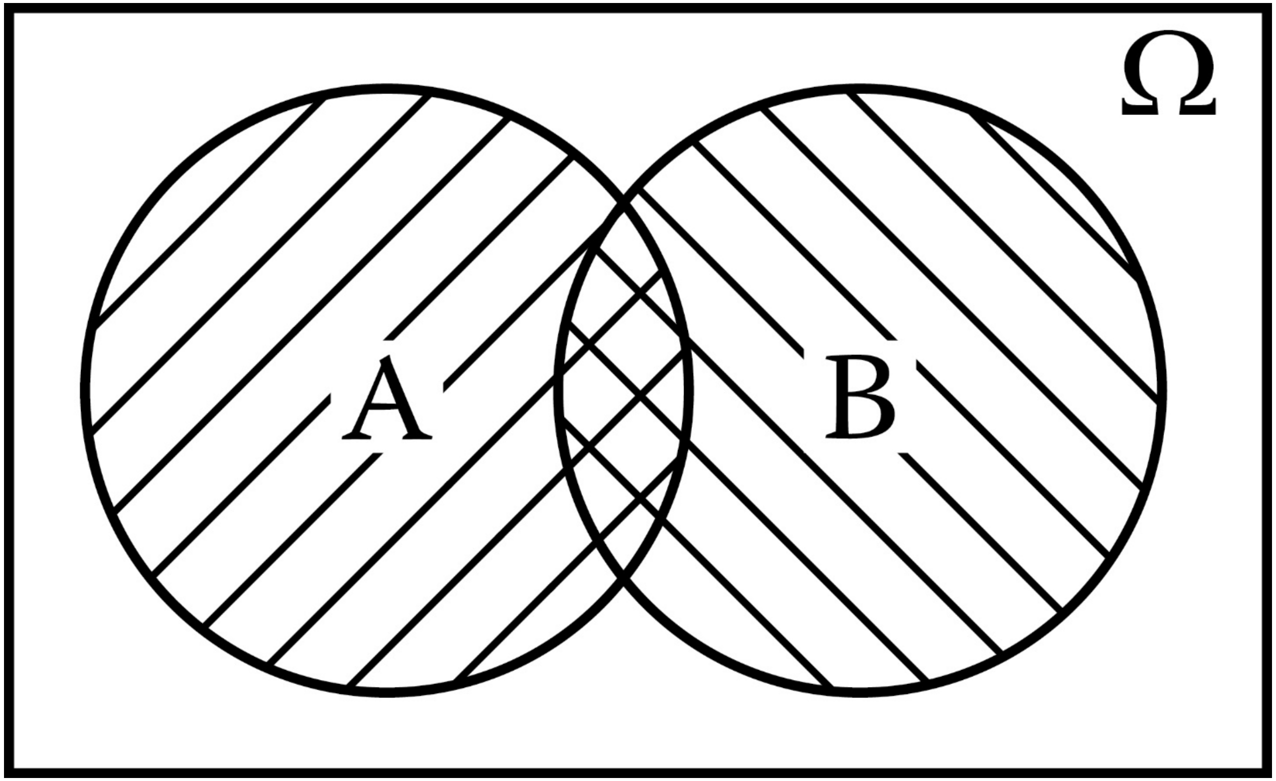
- The **Venn diagram** shows the relationship between sets and subsets.
- Universal set Ω by a rectangle, and its subsets A and B by circles. It can also be seen that B is a subset of A. The relationship between subsets A and B and the universal set can be symbolized by $B \subset A \subset \Omega$ (or *subset equal to*, $\subseteq$).
- In this figure:
  - C has no elements in common with B.
  - A and C have at least one element in common.



Set operation

- The **union** of two sets, A or B, is a set of element that belong to A or B or both.

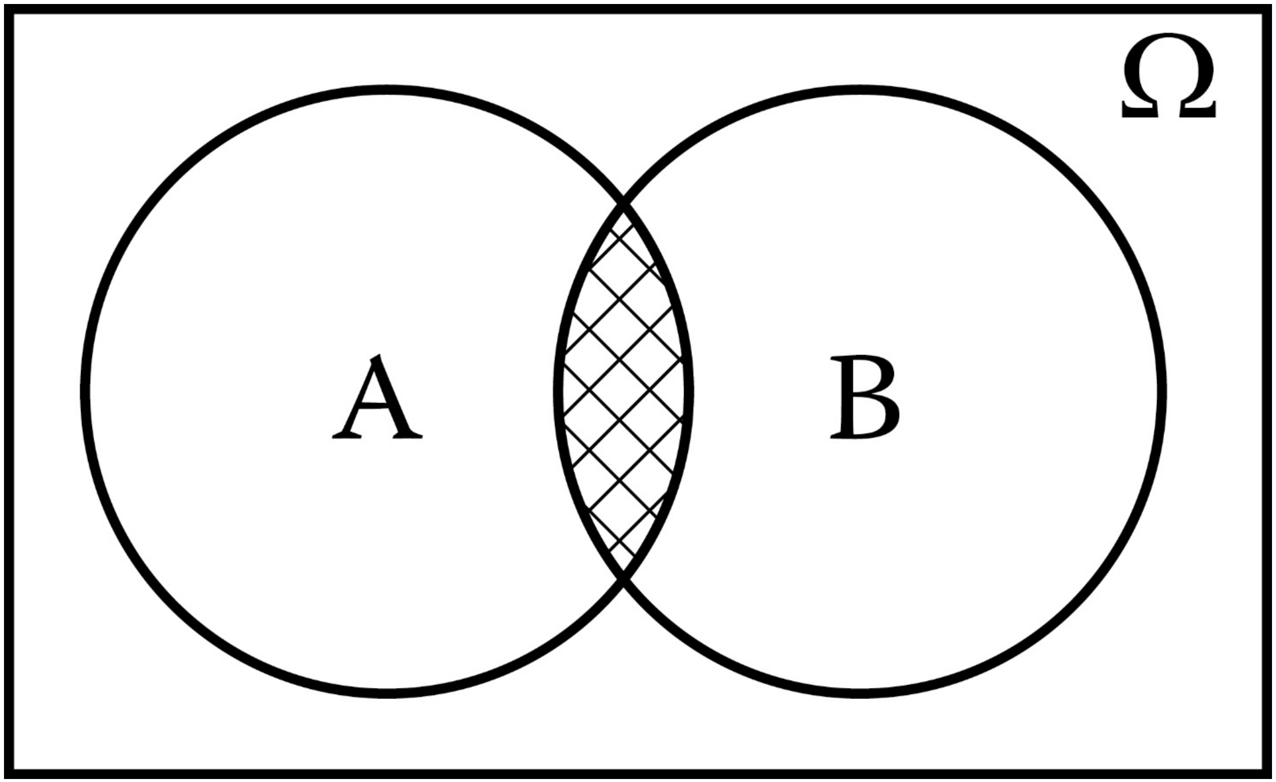
$$(A \cup B) \text{ or } (A + B)$$



Set operation

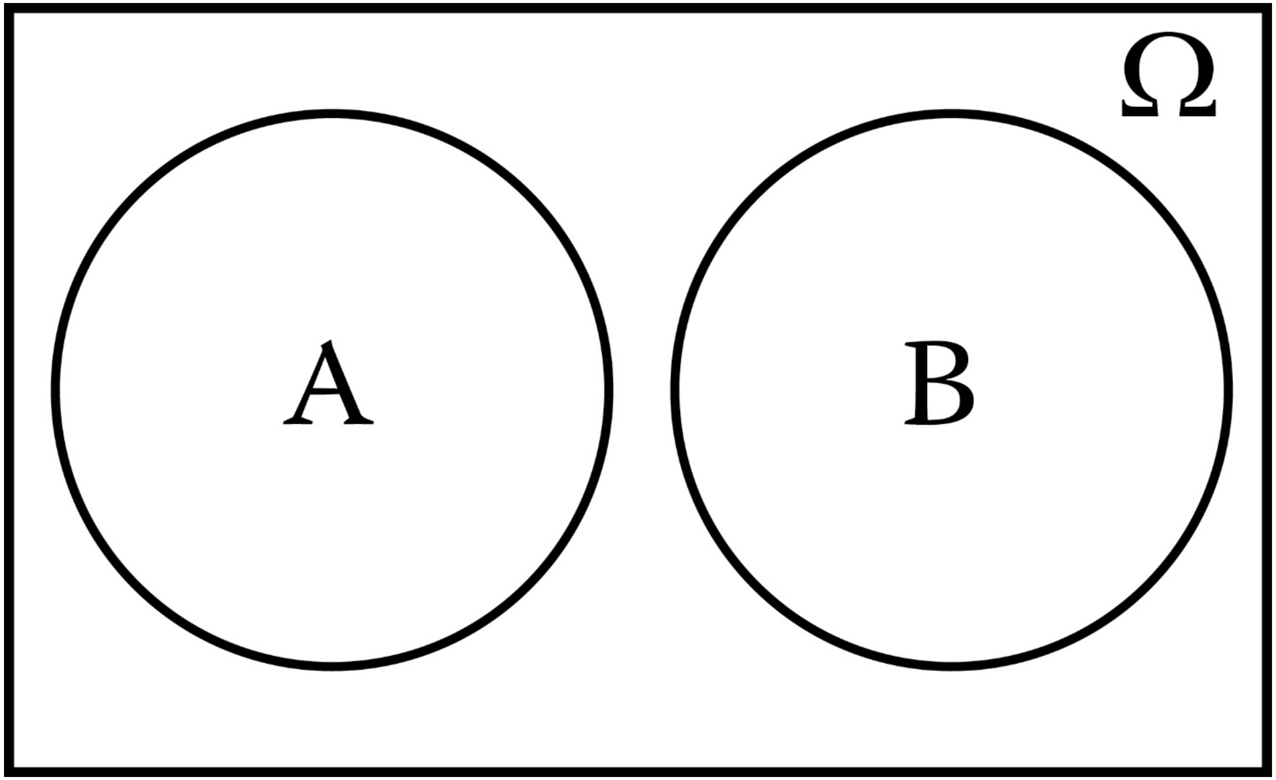
- **Intersection** of two sets called (A and B) is the set of elements belonging to *both* A and B

$$(A \cap B) \text{ or } (A \cdot B)$$



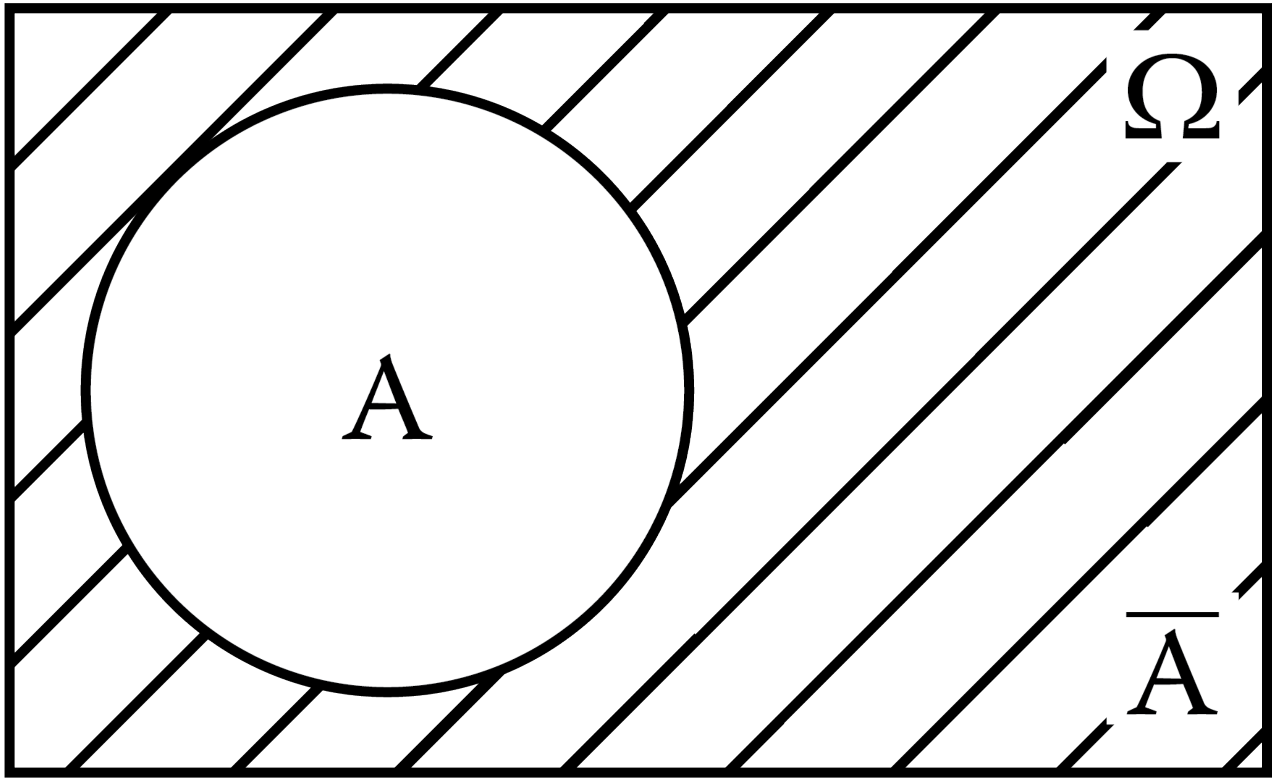
Set operation

- If $A \cap B = \varnothing$, then the sets A and B are **mutually exclusive (or disjoint)** i.e., they have no elements in common.



Set operation

- The **complement** of a subset is all the elements in the universal set that do not belong to subset. The complement of A is $\bar{A}$.



Example: Set operation

- **Example 1**: Consider the discrete set *S* and subset *A*. What is the complement of *A*?

$$S = \{1,2,5,10\} \qquad A = \{2,5\}$$

$Solution$: $\bar{A} = \{1,10\}$

- **Example 2**: Consider the continuous set *S*, which represents lightbulb failure times in our dataset, and subset A, which we will call "early failures." What is the complement of *A*?

$$S = \{t|t > 10\} \qquad A = \{t|100 \geq t > 10\}$$

$Solution$: $\bar{A} = \{t|t > 100\}$

Why we use sets

- We can manipulate interesting subsets of sample spaces or talk about interesting combinations of events.
  - For example, $B = \{t \mid 0 < t < 100\}$ shows a subset of light bulbs that operate for $t > 0$ but fail prior to 100 hours.
  - A = {Failed}
- For example, take the two events: A = pump A has failed (off), and B = pump B has failed (off).
- We might want to know when:
  - $A \cap B$ = Both pumps failed.
  - $A \cap \overline{B}$ = A has failed, B did not fail.
  - $\overline{A} \cap B$= A did not fail, B did fail.
- Boolean algebra is extensively used in reasoning about **events & sets.**

Boolean Algebra Laws

Designation	Mathematical Notation	Engineering Notation
Identity laws	$A \cup \emptyset = A$ $A \cup \Omega = \Omega$ $A \cap \emptyset = \emptyset$ $A \cap \Omega = A$	$A + 0 = A$ $A + 1 = 1$ $A \cdot 0 = 0$ $A \cdot 1 = A$
Idempotent laws	$A \cap A = A$ $A \cup A = A$	$A \cdot A = A$ $A + A = A$
Complement laws	$A \cap \overline{A} = \emptyset$ $A \cup \overline{A} = \Omega$	$A \cdot \overline{A} = 0$ $A + \overline{A} = 1$
Law of Absorption	$A \cup (A \cap B) = A$ $A \cap (A \cup B) = A$	$A + (A \cdot B) = A$ $A \cdot (A + B) = A$
de Morgan's Theorem	$\overline{(\mathrm{A} \cap B)} = \overline{A} \cup \overline{B}$ $\overline{(\mathrm{A} \cup B)} = \overline{A} \cap \overline{B}$	$\overline{(A \cdot B)} = \overline{A} + \overline{B}$ $\overline{(A + B)} = \overline{A} \cdot \overline{B}$
Commutative laws	$A \cap B = B \cap A$ $A \cup B = B \cup A$	$A \cdot B = B \cdot A$ $A + B = B + A$
Associative laws	$A \cup (B \cup C) = (A \cup B) \cup C = A \cup \mathrm{B} \cup C$ $A \cap (B \cap C) = (A \cap B) \cap C = A \cap B \cap C$	$A + (B + C) = (A + B) + C = A + B + C$ $A \cdot (B \cdot C) = (A \cdot B) \cdot C = A \cdot B \cdot C$
Distributive laws	$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	$A \cdot (B + C) = (A \cdot B) + (A \cdot C)$ $A + (B \cdot C) = (A + B) \cdot (A + C)$

Boolean algebra: Example 1

Example: Show that $\overline{(A \cdot B) + \left(A \ \cdot \ \overline{B}\right)} = \ \overline{A}$

(KG reminder: work this out on board immediately; runs too long otherwise).

Boolean algebra: Example 1 solution

Example: Show that $[\overline{(A \cdot B) + \left(A \ \cdot \ \overline{B}\right)} = \ \overline{A}$

$$
\begin{aligned}
\overline{(A \ \cdot B) + \left(A \ \cdot \overline{B}\right)} &= \left(\overline{A \cdot B}\right) \cdot \left(\overline{A \cdot \overline{B}}\right) && \text{(de Morgans)} \\
&= \left(\overline{A} + \overline{B}\right) \cdot \left(\overline{A} + B\right) && \text{(de Morgans)} \\
&= \ \overline{A} + \underbrace{\left(\overline{B} \cdot B\right)} && \text{(Distributive)} \\
&= \overline{A} \quad + \quad 0 && \text{(Complement)} \\
&= \overline{A} && \text{(Identity)}
\end{aligned}
$$

Extra example: Solve this on your own and show that

$$(X \cdot Y) + \left(\overline{X} \ \cdot Z\right) + (Y \cdot Z) = (X \cdot Y) + (\overline{X} \cdot Z)$$

Terminology

- **Probability**: A numerical measure of the chance that an event occurs (or that a hypothesis is true); used to quantitatively express uncertainty about the occurrence of an event
- **Frequency:** The rate of occurrence of events – that is, the number of times an event occurs in a given period of time [or space, or number of trials]
- **Likelihood:** The probability of observing evidence or data when given an event or proposition.
- **Statistics**: Collecting, analyzing, and interpreting data

Probability interpretations used in reliability

- Take an event, X. We want to discuss $\Pr(X = x)$
  - E.g., $\Pr(x) = 0.80$
- Two interpretations are used in aspects of reliability engineering.
  - **Frequentist interpretation: estimate the value of $\Pr(\boldsymbol{x})$ with observed data**
    - Appropriate when the number of observations is large, when the observed data are relevant for (and interchangeable with) the situation at hand.
  - **Subjectivist (Bayesian) interpretation: probability as a degree of belief (state of knowledge)**
    - Allow us to take additional information into account; to deal with rare events, unique systems, one-time events, etc.
- The same laws of probability and Boolean methods apply regardless of interpretation

Philosophical moment….

Probability is not really about numbers;
it is about the structure of reasoning.

Glenn Shafer
Rutgers University

Probability defined

Classical Definition of Probability (sample space partition)

- Associated with any event (outcome) E in a sample space is a probability denoted by:

$$\Pr(E) = \frac{n_E}{n} \text{ (Note that } \Pr(S) = \frac{n}{n} = 1, \Pr(\emptyset) = 0\text{)}$$

- Where:

  n_E = number of times event E occurs in sample space
  n = total number of events in the sample space

Example: All dice outcomes n = 6, All odd dice outcomes $n_E = 3$

Probability defined (cont.)

Frequentist (or Frequency) Definition of Probability

- This interpretation of Pr(E) assumes that the total number of *identical* events in the sample space (n) is unknown. Therefore, the probability of event E may be defined as a limit of n_E/n as n becomes large, i.e. the *proportion (or frequency) of outcomes*:

$$\Pr(E = e) = lim_{n \to \infty} \left(\frac{n_E}{n}\right)$$

- Where:

  n_E = number of times event E occur in sample space
  n = total number of events in the sample space

Example: Number of times a pump is started n = 2000, Number of failed pump starts $n_E = 20$.

Then Pr(E=e) = 20/2000=0.01

Probability defined (cont.)

Subjectivist (or Bayesian) Definition of Probability

- Bayesian probability, also called evidential probability or subjectivist probability), can be assigned to any statement whatsoever, as a way to represent its plausibility, or the degree to which the statement is supported by the available evidence (*degree of belief*).

  - Probability is the **degree of belief** in the truth of a proposition.
  - It is a representation of an individual's state of knowledge
  - The main requirement is **coherence**: One's subjective probability of an event must be consistent with their knowledge (evidence) & the laws of probability
  - ***Assumption***: Any two individuals with the same knowledge, information, and biases will assign the same probability value.

Examples: probability interpretation

Probability Statement	Which probability interpretation could this be? Frequentist or Subjectivist
The probability of flipping a [fair] coin and getting heads is 0.50	
There is an 0.80 probability of rain in College Park tomorrow.	
There is a 0.019 probability that a bolt is defective; the bolt is from a supply of 1000 bolts manufactured today at process plant X.	
The probability that you will receive an A in this class	
The probability of a loss of containment event this year at nuclear power plant X is $1.7*10^{-5}$.	

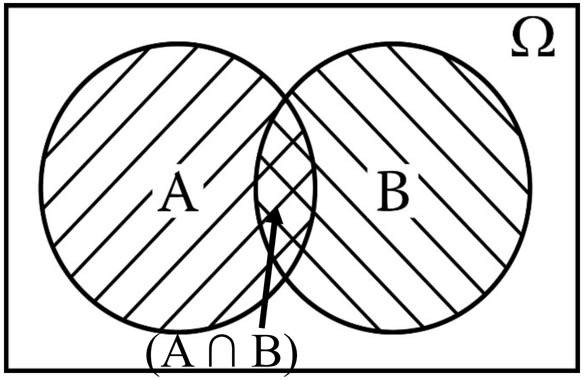
Probability terminology & notation

- **Marginal (unconditional) probability: Pr(A)**
  - The probability of event A occurring
- **Joint probability: Pr(A ∩ B) = Pr(A, B) = Pr(AB)**
  - The probability of events A and B both occurring
- **Conditional probability: Pr(A|B)**
  - The probability of event A occurring given that B has occurred.

$$\mathbf{Pr}(\boldsymbol{A}|\boldsymbol{B}) = \frac{\mathbf{Pr}(\boldsymbol{A} \cap \boldsymbol{B})}{\mathbf{Pr}(\boldsymbol{B})}$$

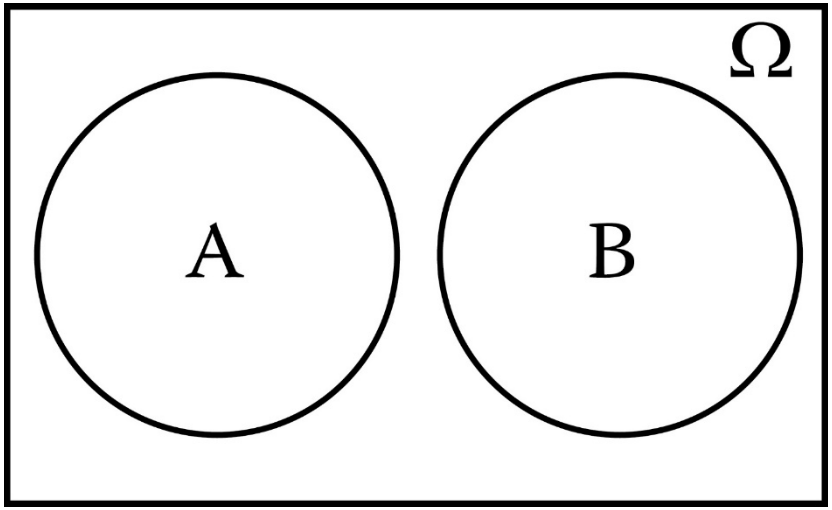
Definition of independent events

- **Definition of independent events:** Two events are independent if the occurrence or nonoccurrence of one does not depend on or change the probability of the occurrence of the other.
  - Mathematically this means: $\Pr(A|B) = \Pr(A)$
    - Which also means that $\Pr(A \cap B) = \Pr(A)\Pr(B)$
  - E.g., A = rolling a 1 on a die (first roll), B = rolling a 1 on the second roll
  - E.g., A = valve A fails to open, B = pump B fails to start (assuming no common cause)



Definition of mutually exclusive (disjoint)

- **Definition of mutually exclusive events:** $A \cap B = 0$
  - Two events are *mutually exclusive* if they can't both happen at the same time
    - Which means that $\Pr(A \cap B) = 0$
  - E.g.: A = rolling an even number on a die, B= rolling an odd number on a die
  - E.g., E1 = Valve E is operational, E2 = Valve E is failed



Axioms of probability

(Kolmogorov 1933):

1. $\Pr(E_i) \geq 0$, for every event E_i

2. $\Pr(\Omega) = 1$

3. $\Pr(E_1 \cup E_2 \cup \ldots) = \Pr(E_1) + \Pr(E_2) + \ldots$,
   when $E_1, E_2, \ldots$ are mutually exclusive
   - Reminder: Mutually exclusive means: i.e., $E_1 \cap E_2 = \emptyset$, no common points exist between $E_1, E_2, \cdots$

Additional implications of the axioms

- $\Pr(\emptyset) = 0$
- $0 \leq \Pr(E_i) \leq 1$
- If $A \subseteq B$, then $\Pr(A) \leq \Pr(B)$
- $\Pr(\bar{E}) = 1 - \Pr(E)$

Chain rule of probability

- The **chain rule of probability (aka multiplication rule):** defines the relationship between joint probability and conditional probability

$$\Pr(E_n \cap E_{n-1} \cap \ \dots \cap E_2 \cap E_1) = \\ Pr(E_n | E_{n-1}, \dots E_2, E_1) \cdot Pr(E_{n-1} | \dots E_2, E_1) \cdot \dots \cdot Pr(E_2 | E_1) \cdot Pr(E_1)$$

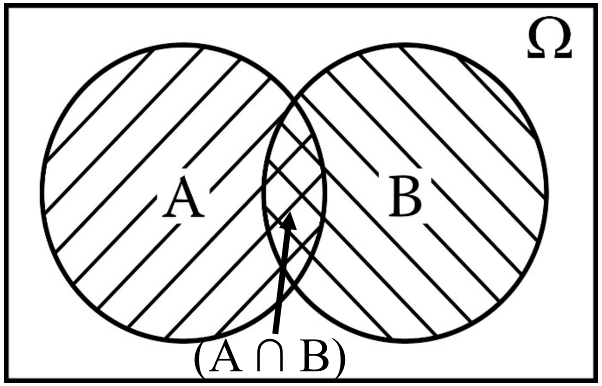
- If all events are **independent** (that is, $E_n \perp E_{n-1} \perp \ \dots \ \perp E_2 \perp E_1$), this simplifies to:

$$Pr(E_n \cap E_{n-1} \cap \ \dots \cap E_2 \cap E_1) = \prod_{i=1}^{n} Pr(E_i) \\ = Pr(E_n) \cdot Pr(E_{n-1}) \cdot \dots \cdot Pr(E_2) \cdot Pr(E_1)$$

Addition law of probability

The **addition law of probability** (**inclusion-exclusion principle**)

If there are common element between A and B (i.e., they are not mutually exclusive, $A \cap B \neq \varnothing$) then:
$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$.



If all E_i events are independent this is written in compact form as:

$$\Pr(E_1 \cup E_2 \cup \cdots \cup E_n) = 1 - \prod_{i-1}^{n}(1 - \Pr(E_i))$$

Law of total probability

- The **law of total probability** defines the relationship between joint and marginal distributions:

$$\Pr(A) = \sum_{i=1} \Pr(A \cap B_i) = \sum_{i=1} \Pr(A|B_i)\Pr(B_i)$$

- For example, to marginalize out a binary variable B:

$$\Pr(A) = \Pr(A \cap B) + \Pr(A \cap \overline{B})$$
$$\Pr(A) = \Pr(A|B)\Pr(B) + \Pr\left(A\middle|\overline{B}\right)\Pr(\overline{B})$$

Example

- **Example:** You receive bolts from two *independent* suppliers. The probability of the event D_i, "selecting a defective bolt from supplier *i*" is characterized by:

$$\Pr(D_1) = 0.05,\ \ \Pr(D_2) = 0.07$$

- If one bolt is selected from each supply, find:
    a) The probability that both selected bolts are defective?

    b) The probability that at least one is defective?

Hint: start by writing the LHS side of the questions in set & probability notation.

Example Solution

Example: Given $\Pr(D_1) = 0.05$, $\Pr(D_2) = 0.07$, and $D_1 \perp D_2$. One bolt is selected from each supplier.

a) The probability that both selected bolts are defective, since the bolts are independent, is:

$$\boldsymbol{Pr(D_1 \cap D_2)} = \Pr(D_1)\Pr(D_2) = (0.07)(0.05) = \mathbf{0.0035}$$

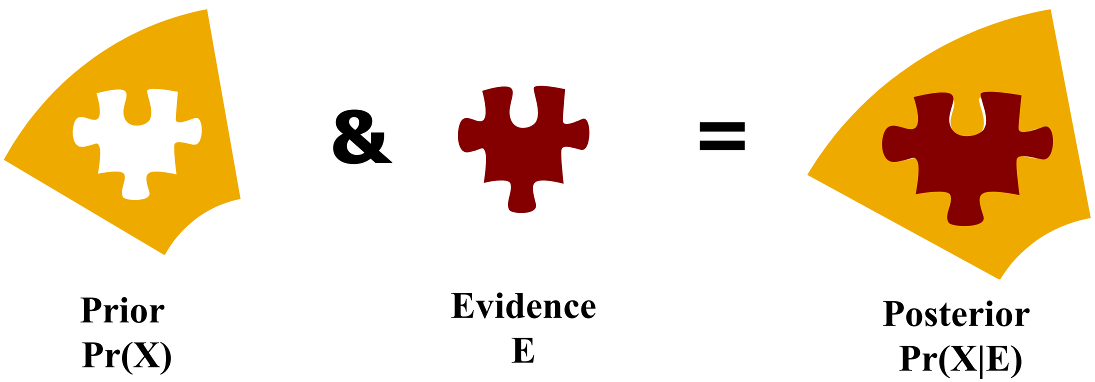
b) The probability that at least one bolt is defective is:

$$\boldsymbol{Pr(D_1 \cup D_2)} = \Pr(D_1) + \Pr(D_2) - \Pr(D_1 \cap D_2)$$

$$= 0.05 + 0.07 - 0.0035 = \mathbf{0.1165}$$

Bayes' Theorem: Conceptual

- Probability is used as the foundation for *reasoning*
- Provides a framework for updating beliefs (probabilities) in light of new evidence.



Bayes' Theorem (discrete)

- If A_j is an event of interest, and if E is any other event (or set of evidence) such that Pr(E) > 0, then:

$$\mathbf{Pr}\left(\boldsymbol{A_j}\middle|\boldsymbol{E}\right) = \frac{\mathbf{Pr}(\mathbf{A}_{\boldsymbol{j}})\mathbf{Pr}\left(\boldsymbol{E}\middle|\boldsymbol{A_j}\right)}{\mathbf{Pr}(\mathbf{E})}$$

  - $\Pr\left(\mathrm{A_j}\middle|\mathrm{E}\right)$ is called **posterior (or a posteriori) probability of $\mathbf{A_j}$** (given evidence E)
  - $\Pr(\mathrm{A_j})$ is called **prior (or a priori) probability of $\mathbf{A_j}$**
  - $\Pr(\mathrm{E}) = \sum_{\mathrm{i}=1}^{\mathrm{n}} \Pr(\mathrm{A_i})\Pr(\mathrm{E}|A_i)$ from the law of total probability.
  - $\Pr\left(E\middle|A_j\right)/\Pr(E)$ **is the relative likelihood** (the model by which the probability is revised)

Bayes' Theorem (continuous)

- For **continuous probability functions**, Bayes' Theorem is given:

$$f(\theta|E) = \frac{\Pr(E|\theta)\Pr(\theta)}{\Pr(E)} = \frac{f(\theta)f(E|\theta)}{f(E)}$$

$$f(E) = \int f(E|\theta)f(\theta)d\theta \longleftarrow \text{marginal pdf of } E$$

Example: Events, probability & Bayes' Theorem

- Tractors made by a company are from three assembly lines: Red, White, and Blue. (R, W, B)
- Suppose 48% of the company's tractors are made on the Red line and 31% are made on the Blue line.
- Let D be the event that the new tractor is defective (i.e., it wont start)
- The probability that a tractor will not start when it rolls off of a line are 6%, 11%, and 8% given that it comes from Red, White, and Blue, respectively.

a) Translate the problem statement & questions into probability notation.

b) What is the probability that a randomly selected tractor is defective?

c) What is the probability that a tractor came from the red line given that it is defective?

Solution part a: Define events

Events:

- Let D be the event that the tractor won't start
- Let R be the event that the tractor was made by the red assembly
- Let W be the event that the tractor was made by the white assembly
- Let B be the event that the tractor was made by the blue assembly

$$\Pr(\mathrm{R}) = 0.48$$

$$\Pr(\mathrm{W}) = 0.21$$

$$\Pr(\mathrm{B}) = 0.31$$

$$\Pr(\mathrm{D}|\mathrm{R}) = 0.06$$

$$\Pr(\mathrm{D}|\mathrm{W}) = 0.11$$

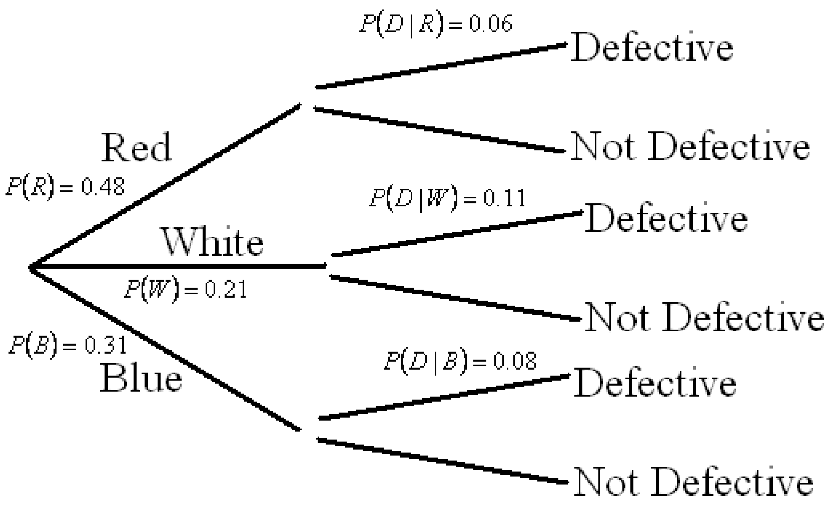
$$\Pr(\mathrm{D}|\mathrm{B}) = 0.08$$

- Our problem:

  b) What is the probability that a randomly selected tractor is defective? In other words: **Find $\mathbf{Pr(D)}$**.

  c) What is the probability that a tractor came from the red line given that it was defective? **Find $\boldsymbol{Pr(R|D)}$**.

Solution part b & c: Tree diagram



$\Pr(R \cap D)$?

$\Pr(W \cap D)$?

$\Pr(B \cap D)$?

$$\Pr(R \cap D) = \Pr(D|R) \Pr(R)$$

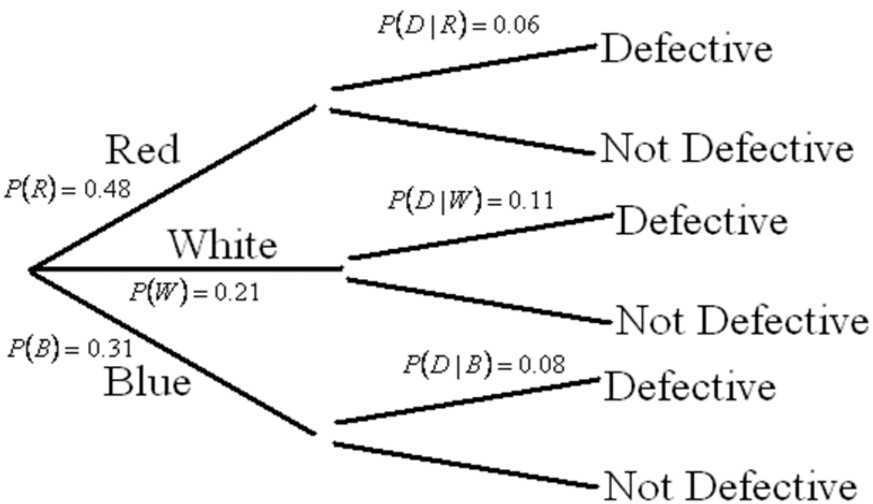
$$\Pr(W \cap D) = \Pr(D|W) \Pr(W)$$

$$\Pr(B \cap D) = \Pr(D|B)\Pr(B)$$

- Because each of these events represents an instance where a tractor is defective, to find the total probability that a tractor is defective, regardless of who made it:

$$\Pr(D) = \Pr(D|R) \Pr(R) + \Pr(D|W) \Pr(W) + \Pr(D|B)\Pr(B)$$

Solution parts b & c



$\Pr(R \cap D)$?
$\Pr(W \cap D)$?
$\Pr(B \cap D)$?

$$\Pr(R \cap D) = \Pr(D|R)\Pr(R)$$
$$\Pr(W \cap D) = \Pr(D|W)\Pr(W)$$
$$\Pr(B \cap D) = \Pr(D|B)\Pr(B)$$

- **Solution:**

$$\Pr(\mathrm{D}) = \Pr(\mathrm{D}|\mathrm{R})\Pr(\mathrm{R}) + \Pr(\mathrm{D}|\mathrm{W})\Pr(\mathrm{W}) + \Pr(\mathrm{D}|\mathrm{B})\Pr(\mathrm{B})$$

$$\Pr(D) = (0.06)(0.48) + (0.11)(0.21) + (0.08)(0.31)$$

$$\boxed{\mathbf{Pr}(\boldsymbol{D}) = \mathbf{0.077}}$$

$$\mathbf{Pr}(\boldsymbol{R}|\boldsymbol{D}) = \frac{Pr(R \cap D)}{Pr(D)} = \mathbf{0.374}$$

Probabilities—where do they come from? (1)

- **The probability of an event is not directly observable**
- **Probability models help translate data into probabilities.**
  - We can observe if events (e.g., failures) occur (about x failures out of n trials), and the times of failures $(t_1, t_2, t_3, \ldots)$
  - We can count combinations of items or options
  - We can collect measurements, run large sets of simulation models, etc.
  - We can ask experts to assign probabilities ("implicit models")

Probabilities – where do they come from? (2)

- Random Variables may be **Discrete** or **Continuous**.
- Similarly, probability distributions representing the Random Variables are also discrete or continuous
  - Binary outcomes (e.g., success/fail, on/off)
  - Discrete data (e.g., countable items)
  - Continuous quantities (e.g., measured variables)

Probability distributions

- **Definition**: A valid **probability distribution** model for a random variable is a function which assigns a total probability of 1.0 to the set of all points in the sample space S.

$$S = \{x_1, x_2, \ldots, x_k\} \text{ where } k \text{ is either finite or infinite}$$

The probability model (**probability density function**) is then one for which:

$$\mathrm{f}(x_i) \geq 0 \text{ for } i = 1,2,\ldots,k \quad \text{or } -\infty < x < \infty$$

and

$$\sum_{i=1}^{k} f(x_i) = 1 \quad \text{or } \int_{-\infty}^{\infty} f(x)dx = 1 \text{ (i.e., the area under the curve is 1)}$$

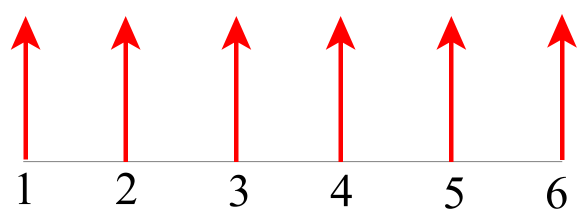
Discrete uniform distribution

- For a sample space $S = \{x_1, x_2, \dots, x_k\}$ with k equally likely outcome; that is, the probability of each outcome is:

$$\mathrm{f(x_i)} = \Pr(X = x_i) = \frac{1}{k}, \qquad i = 1, 2, \dots, k$$

- Example: Take a die. The probability that any one of the die options $\Omega = \{1,2, \dots, 6\}$ will be rolled is:

$$\Pr(x_i) = \frac{1}{6}$$



Probability density function (PDF)

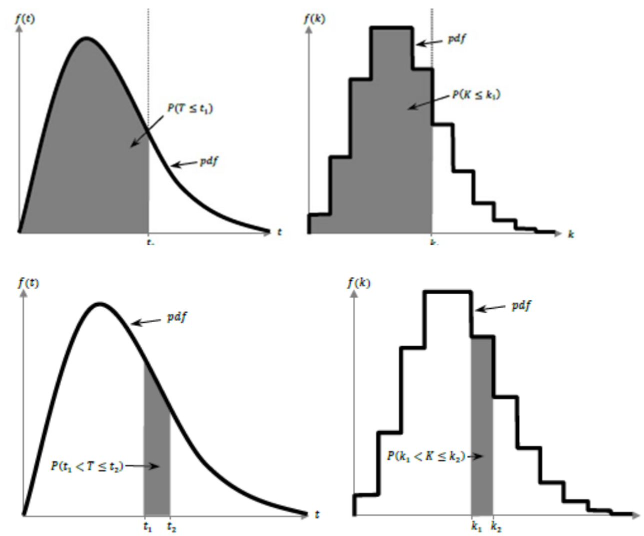
- The function **f(x)** is the **probability density function** (pdf).
  - For discrete distributions, **f(x)** is the **probability mass function** (pmf).
- The pdf/pmf provides point probabilities (discrete r.v.) or point densities (continuous r.v.).
- For a random discrete variable X:

$$\boldsymbol{f(x) = \mathrm{Pr}(X = x)}$$

- For a continuous r.v. X, we use the pdf to solve for the **cdf**:

$$\boldsymbol{Pr(X \le x_i) = \int_{-\infty}^{x_i} f(x)dx = F(x_i)}$$

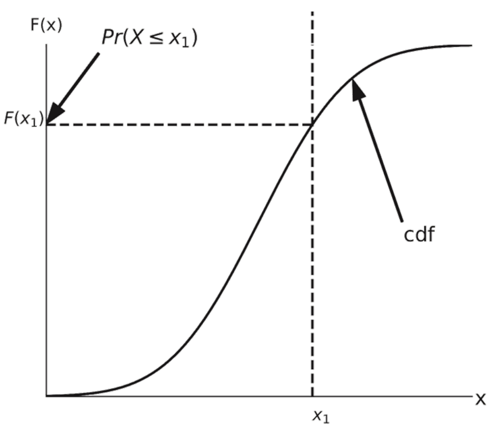
$$\boldsymbol{\mathrm{Pr}(a \le X \le b) = \int_{a}^{b} f(x)dx = F(b) - F(a)}$$

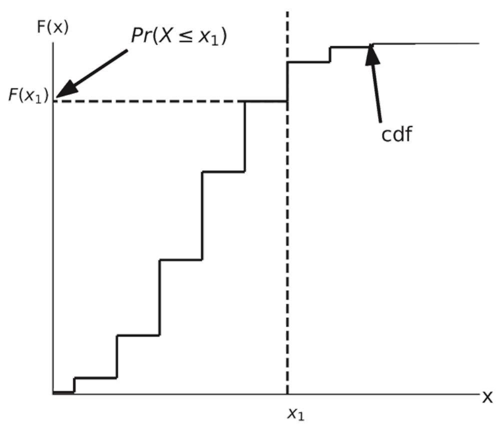


Cumulative distribution function (cdf)

- **Definition**: The **cumulative distribution (density) function** (or **cdf**) for an r.v. X is defined as:

$$F(x) = \begin{cases} \Pr(X \leq x) = \sum_{x_i \leq x} \Pr(x_i) & \text{for discrete r.v.}\, X \\ \Pr(X \leq x) = \int_{-\infty}^{x} f(t)dt & \text{for continuous r.v.}\, X \end{cases}$$



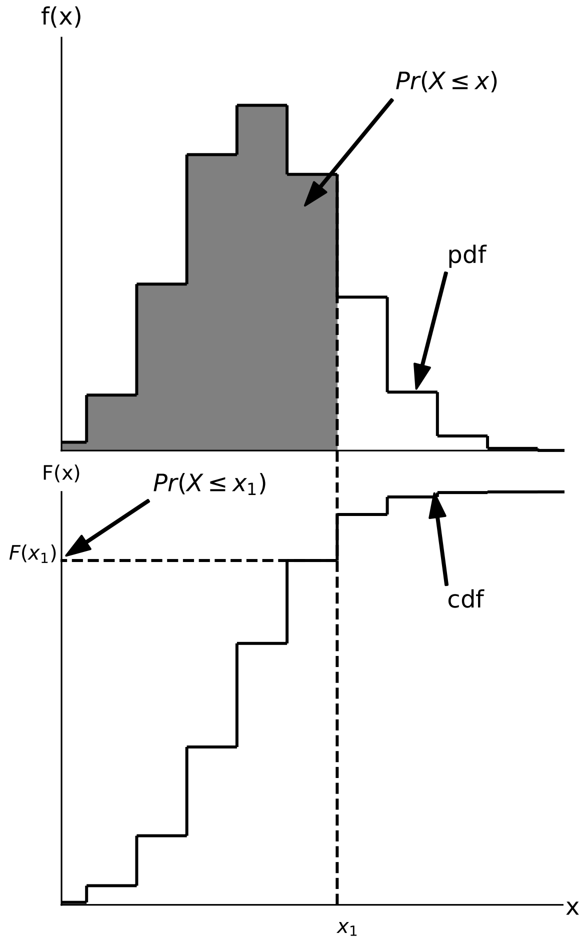


Note:

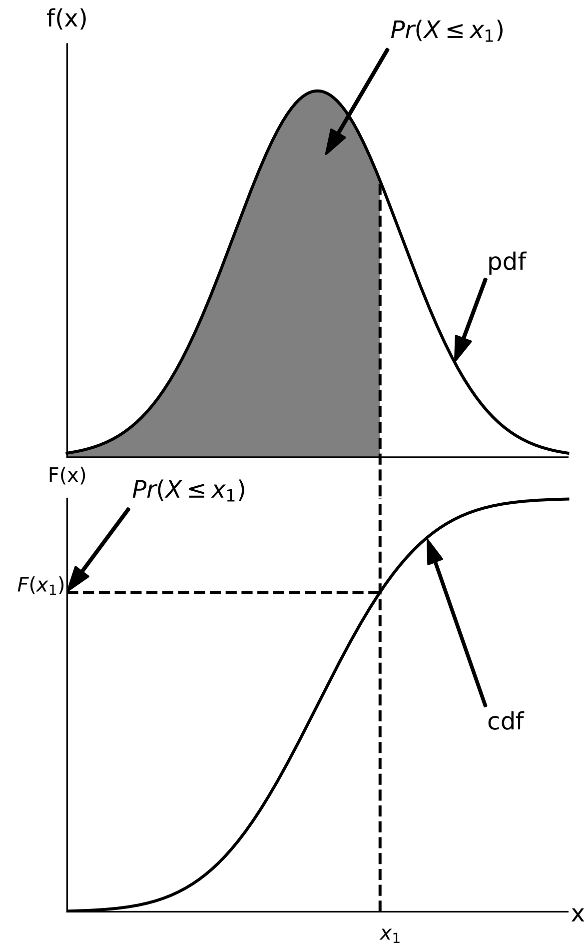
$$f(x) = \frac{dF(x)}{dx} = F'(x)$$

$$\sum_{x \leq \infty} \Pr(x) = 1 \text{ and } \int_{-\infty}^{\infty} f(x)dx = 1$$

The pdf and cdf for discrete (L) and continuous (R) distributions



Discrete pdf (top) and cdf (bottom).

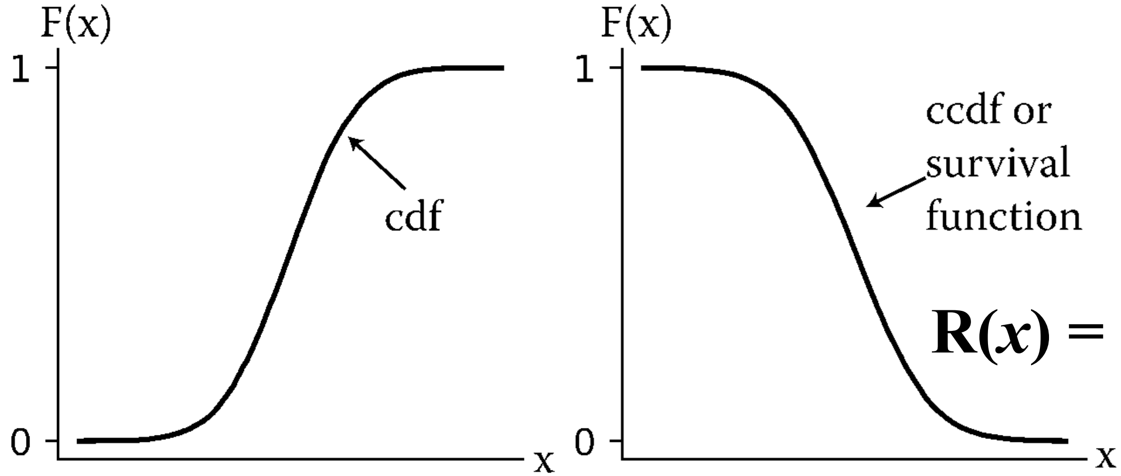


Continuous pdf (top) and cdf (bottom).

Cumulative probability models

- The **complementary cumulative density functions** (or **ccdf**) is also useful especially in instances where a product can fail or survive. This is expressed as:

$$\boldsymbol{R(t) = 1 - F(x)} = \begin{cases} \Pr(X > x) = 1 - \sum_{x_i \le x} \Pr(x_i) & \text{for discrete r.v. } X \\ \int_x^{\infty} f(t)dt & \text{for continuous r.v. } X \end{cases}$$



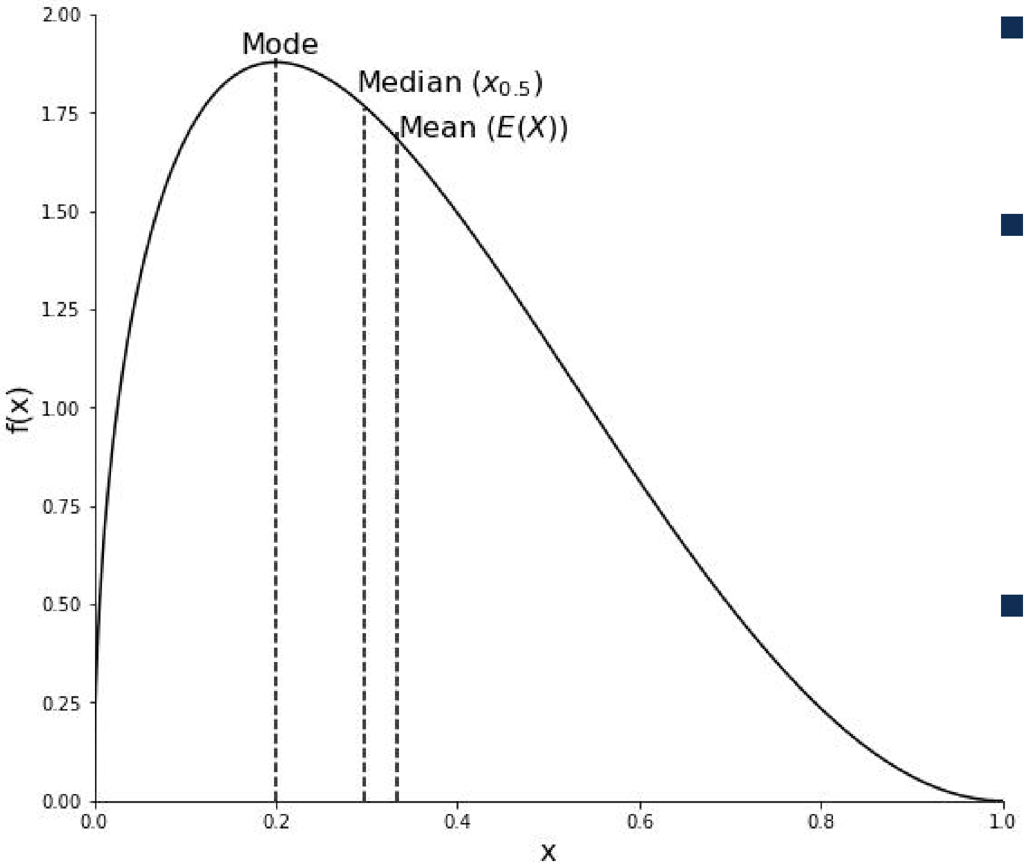
R(*x*) = 1 – F(*x*)

Probability models

- Many parametric probability distributions exist – they are families of similarly-shaped distributions which vary based on the defined parameters.

- Many types of parameters in probability distributions:
  - **Location** – shifts the location around x (or t) axis
  - **Scale** (or dispersion) – spread (stretch/shrink)
    - **Rate** (The reciprocal of scale, i.e., 1/scale)
  - **Shape** – defines distribution shape.

Describing distributions: Measures of central tendency



- **Mean**: The expected value of the r.v.: E(X)
- **Median** ($x_{0.5}$): the point where the cdf is equal to 0.5.
  - $x_{0.5} = \mathbf{F^{-1}(0.5)}$
- **Mode**: The highest point of a PDF (the value of the r.v. which has the highest probability of occurrence)

Expected value of X

- For a discrete r.v. X with sample space $\{x_1, x_2, \dots, x_k\}$, and probabilities $\Pr(x_i)$, the expected value of X, E(X) is:

$$E[X] = \sum_{i=1} x_i \, f(x_i)$$

- Similarly for continuous distribution with a pdf of f(x)

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

(Riemann integral)

Expected value: generalized

- The same concepts apply to obtain the expected value of any real-value function of X:
  - For a discrete r.v. X with sample space $\{x_1, x_2, \dots, x_k\}$, if $g(x)$ is a real-valued function of X then expected value of $g(x)$ is defined as:

$$E[g(X)] = \sum_{i=1}^{k} g(x_i) \cdot \Pr(x_i)$$

  - Similarly, for continuous distribution with a pdf of f(x):

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) \cdot f(x)\, dx$$

- And for joint distributions:

$$\mathrm{E}[g(x_1, x_2, \dots, x_n)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, x_2, \dots, x_n) f(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n$$

Additional ways to describe distributions

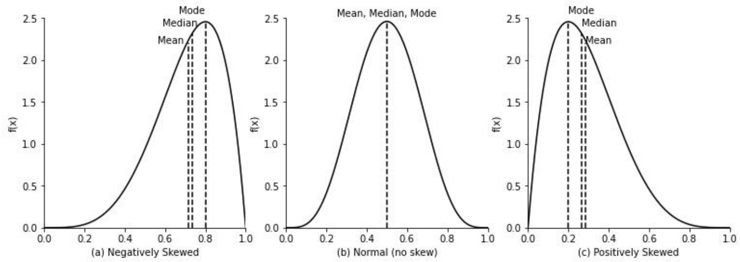
- Some expectation values are given special names and tell us about the shape of the distribution or data.

  $\mu_k = E[X^k] \quad k = 1,2,\ldots$ k^{th} moment about the origin

  $\mu_k = E[(X-\mu_1)^k], \quad k = 2,3,\ldots$ k^{th} moment about the mean (central moments)

  $\mu'_k = E\left[\left(\frac{X-\mu_1}{\sigma}\right)^k\right], \quad k = 2,3,\ldots$ k^{th} standardized moment

- ***We primarily use:***
  - **Mean** (1st raw moment, μ_1), E[X]
  - **Variance** (2nd central moment, μ_2); $\mathrm{Var}(X) = \sigma^2 = E[(X-\mu_1)^2]$
  - **Skewness** (3rd standardized moment, μ'_3); Skew(X) is a measure of the horizontal symmetry (or asymmetry) of the distribution.



  - **Kurtosis** (4th standardized moment, μ'_4) Kurt(X), k, is the measure of whether the distribution is peaked or flat

Example: Calculating expected values

- **Example**: Find E[t] and E[t^2] for t, an exponentially distributed r.v. with $f(x) = \lambda e^{-\lambda t}$
- $E[t] = \int_0^\infty t\lambda e^{-\lambda t} dt$ **Use integration by parts**

$$= -te^{-\lambda t} - \int_0^\infty -e^{-\lambda t} dt$$

$$= -te^{-\lambda t} \Big|_0^\infty - \frac{e^{-\lambda t}}{\lambda} \Big|_0^\infty$$

$$= (0 - 0) - (0 - \tfrac{1}{\lambda})$$

$$= \frac{1}{\lambda}$$

- Similarly for g(x) = t^2:

$$E[t^2] = \int_0^\infty t^2 \lambda e^{-\lambda t} dt = \frac{2}{\lambda^2}$$

Covariance and correlation

- **Definition**: We define **covariance**, as measure of dependence between two random variables. (how much they vary together)

$$\begin{aligned}\mathrm{Cov}(X,Y) &= \mathrm{E}\{[X-\mathrm{E}(X)][Y-\mathrm{E}(Y)]\} \\ &= \mathrm{E}(XY)-\mathrm{E}(X)\mathrm{E}(Y)\end{aligned}$$

- Covariance is difficult to interpret, so we normalize it into **correlation**, defined between X_1 and X_2 as:

$$\rho(X,Y)=\frac{\mathrm{Cov}(X,Y)}{\sigma(x)\sigma(y)}$$

  - $\pm(0,1)$ correlation coefficient indicates strength of relationship; sign indicates direction of relationship. 0 indicates independence.

Combinations: The binomial coefficient

- A specialized way of counting is "Combinations" defined as combination of ***r*** items that can be selected from ***n*** distinct items without replacement (**order not important**)

$$\boxed{C_r^n = \binom{n}{r} = \frac{n!}{r!\,(n-r)!} \qquad r = 1,2,\dots,n}$$

- Definition: $1! = 1, \quad 0! = 1$

- **Example:** From 20 light bulbs how many samples of size 4 can be selected?

$$C_4^{20} = \binom{20}{4} = \frac{20!}{4!\,(20-4)!} = 4845$$

Binomial distribution, *X~binom(n,p)*

- Applies to a series of n independent trials where the random variable X may take binary values (e.g., success or failure, on or off, heads or tails) with probability p.

- The probability that outcome #1 occurs exactly x times out of n trials is given by:

$$f(x|n,p) = \begin{cases} \Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x} & x = 0, 1, 2, \ldots, n \\ 0 & otherwise \end{cases}$$

where,
  - p =probability of outcome #1
  - n =size of the sample space
  - In Excel: binom.dist(x,n, p, false)
  - In Matlab: Binopdf(x,n,p)

Binomial CDF

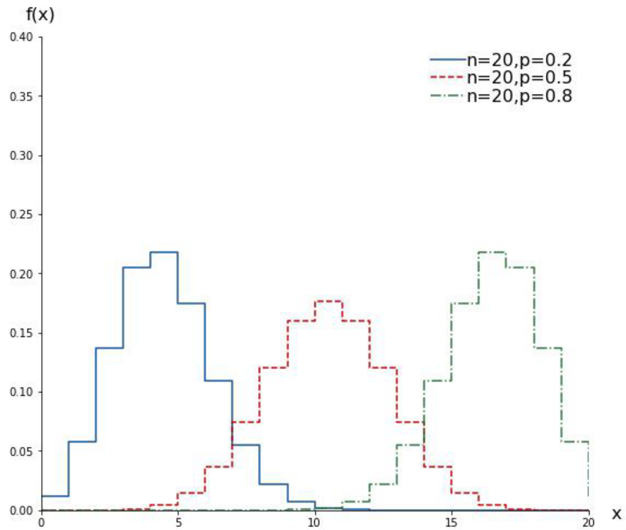
- The **Binomial cdf is**

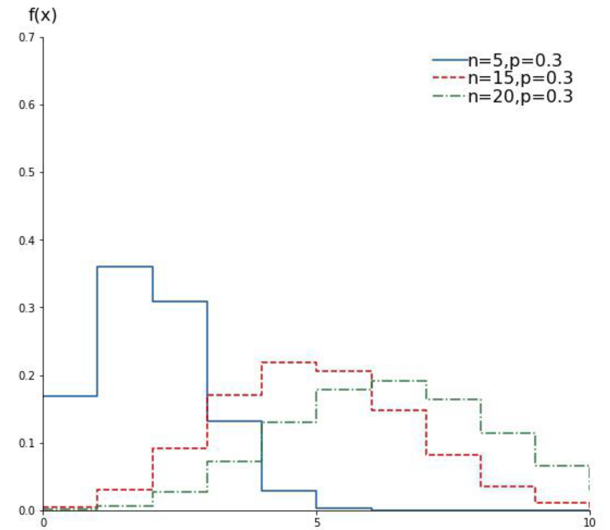
$$F(x|n,p) = \begin{cases} 0 & x < 0 \\ \displaystyle\sum_{i=0}^{x} \binom{n}{i} p^i (1-p)^{n-i} & x = 0,1,2,\ldots,n \\ 1 & x \geq n \end{cases}$$

In Excel: binom.dist(x,n, p, true)
In Matlab: binocdf
Excel is always: (false for PDF, true for CDF)

Binomial distribution (cont.)

Binomial pdf:

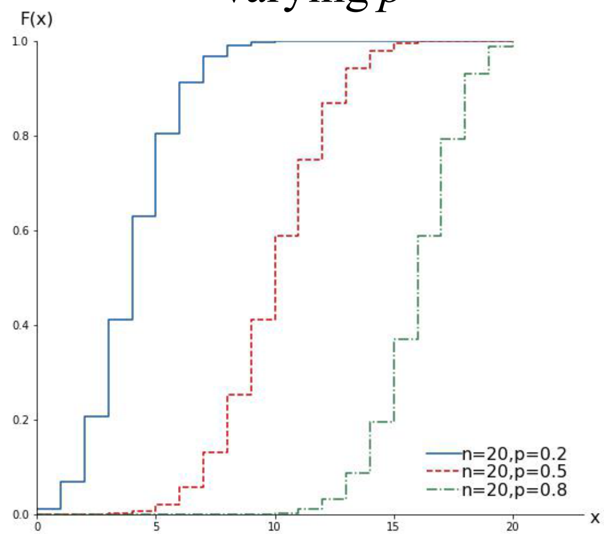


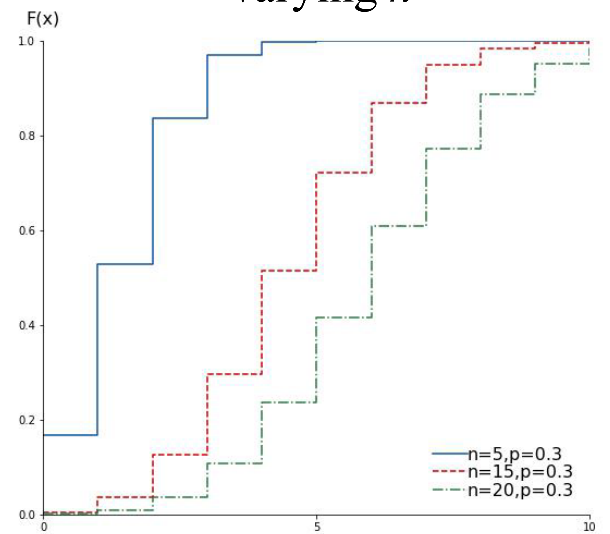


Varying p

Varying n

Binomial cdf:





Binomial distribution

- Example: Assume p, the probability of a light bulb surviving until the end of a specified mission, is given as: $p = 0.9$. For a set of 4 randomly selected light bulbs, calculate probability of:
- A) Exactly 0, 1, 2, 3, or 4 survivals.
- B) 2 or more survivals.

Binomial distribution

- **Solution**: Assume p =probability of survival (success) of light bulb in a mission. That is $p = 0.9$. For $n = 4$ light bulbs, calculate probability of X=0, 1, 2, 3, or 4 survivals.

A)

x	$\Pr(x)$	$F(x) = \Pr(X \leq x)$
0	$\binom{4}{0}(0.9)^0(0.1)^4 = \mathbf{0.0001}$	0.0001
1	$\binom{4}{1}(0.9)^1(0.1)^3 = \mathbf{0.0036}$	0.0001+0.0036=0.0037
2	$\binom{4}{2}(0.9)^2(0.1)^2 = \mathbf{0.0486}$	0.0001+0.0036+0.0486=0.0523
3	$\binom{4}{3}(0.9)^3(0.1)^1 = \mathbf{0.2916}$	0.3439
4	$\binom{4}{4}(0.9)^4(0.1)^0 = \mathbf{0.6561}$	1.0
Total	1.0000	

B) $\mathbf{Pr(X \geq 2)}$ =Pr(2)+Pr(3)+Pr(4) = **0.9963**

$= 1 - \Pr(X<2)$=1-.0037=0.9963

Poisson distribution, *X~poiss(μ)*

- **Definition**: The **Poisson Distribution** expresses probability of a given number of events (*X*) in a time or space, for events that occurs with a constant rate (or intensity), λ & independently of time since last event.
- **Example**: One can think of *X* as the number of electronic components that fail per unit of time or number of telephone calls coming in per unit of time.

$$f(x|\mu) = \Pr(X = x) = \begin{cases} \dfrac{\mu^x e^{-\mu}}{x!} & \mu > 0, x = 0,1,2,\ldots \\ 0 & otherwise \end{cases}$$

  - x – The number of times the event takes place
  - $\mu = \lambda t$ is the rate parameter. It's also the expected (mean) number of occurrences in a fixed time period, t.
    - λ – The (constant) event rate or intensity (Units: 1/time)
    - t – The unit of time or space (Units: time)
  - In Excel: poisson.dist(x, μ, false)

Poisson CDF

- The **Poisson cdf is**

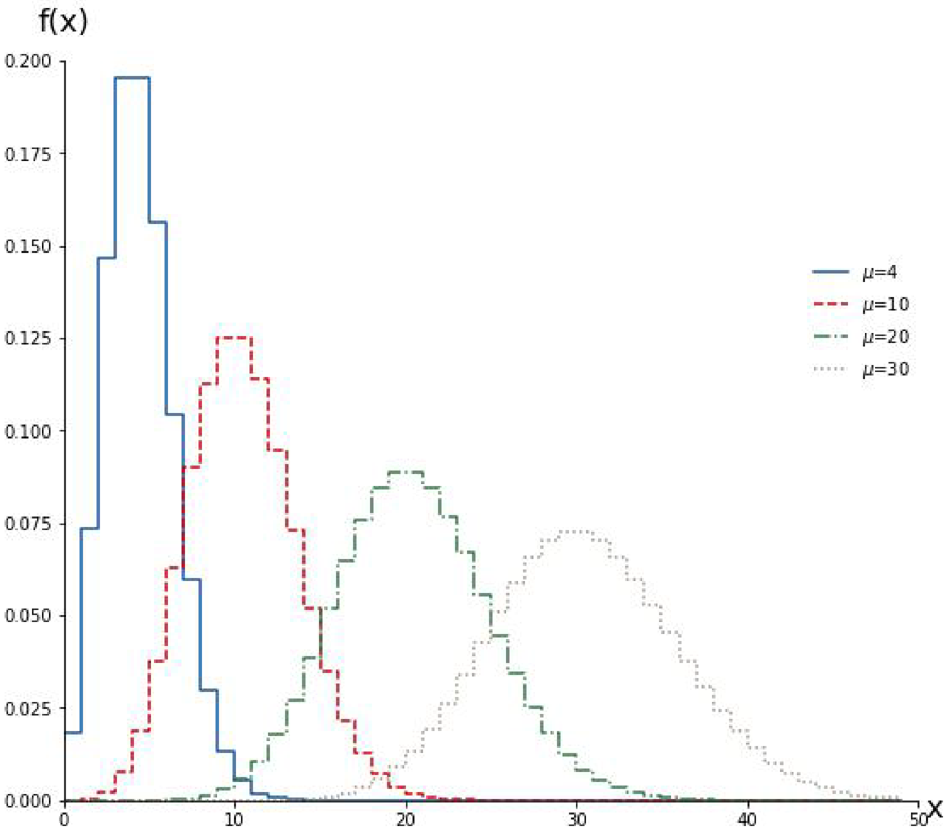
$$F(x|\mu) = \begin{cases} e^{-\mu} \sum_{i=0}^{x} \frac{\mu^i}{i!} & \mu > 0, x = 0,1,2,\ldots \\ 0 & otherwise \end{cases}$$

Matlab function poisscdf(x, μ) to find it.

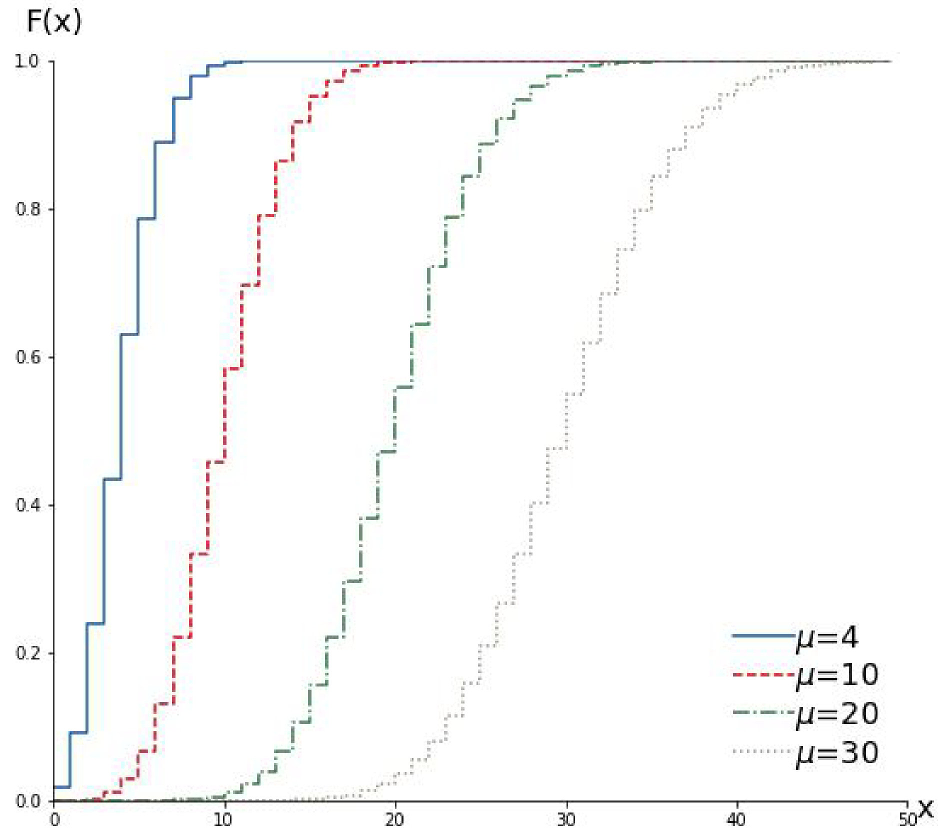
Note: If X is a binomially distributed variable with small p (say $p \leq 0.1$) the Poisson distribution can be used with $\mu = np$.

Poisson distribution (cont.)

pdf



cdf



Poisson distribution: Example 1

- **Example:** Components needed for operation of a machine fail according to a Poisson process with rate of λ = 1 failure/week. What is the probability that exactly 1 component fails during a 4-week period?
  - The expected number of events in 4 weeks is calculates to:

$$\mu = \left(\frac{1}{\text{week}}\right) 4 \text{ weeks} = 4$$

  - The probability for $X = 1$ component failing when $\mu = 4$ is:

$$\Pr(X = 1) = \frac{4^1 e^{-4}}{1!} = 0.0733$$

- **Example part 2:** For this machine, one spare component is available. What is the probability that the machine works for a given 6 weeks?

Poisson distribution: Example 1

- **Example (Part 2)**: For that machine, one spare component is available. What is the probability that the machine works for a given 6 weeks?
  - Here you are solving for X = number of failures

$$\Pr(X \leq 1) = \Pr(X = 0 \ \cup \ X = 1)$$
$$= \Pr(X = 0) + \Pr(X = 1)$$
$$\text{With } \mu = 6$$
$$= e^{-6} + 6e^{-6} = \mathbf{0.0174}$$

Poisson distribution: Example 2 (after class)

Example: Cars typically arrive at a certain stop sign at a rate λ of 3 cars per minute.

- What is the probability that no car arrives in a 10-minute interval?

- What is the probability that more than 20 cars arrive in a 10-minute interval?

Poisson distribution: Example 2

Example: Cars typically arrive at a certain stop sign at a rate λ of 3 cars per minute.

- What is the probability that no car arrives in a 10-minute interval?

$$\mu = \lambda t = \left(3\,\frac{\text{cars}}{\text{minute}}\right) 10 \text{ minutes} = 30 \text{ cars}$$

$$\Pr(x = 0 \text{ cars}|\lambda\text{t} = 30) = \frac{30^0 e^{-30}}{0!} = \boldsymbol{e^{-30}} = \mathbf{9.36 \times 10^{-14}}$$

Excel (poisson.dist(0,30,false))

- What is the probability that more than 20 cars arrive in a 10-minute interval?

$$\Pr(x > 20 \text{ cars}|\lambda\text{t} = 30) = 1 - \Pr(x \leq 20)$$
$$= 1 - \Pr(x = 20) - \Pr(x = 19) - \cdots - \Pr(x = 0)$$

Use Matlab poisscdf(x, μ) or Excel (poisson.dist(20,30,true) to find

$$Pr(X \leq 20) = 0.0352$$

$$1 - Pr(X \leq 20) = \Pr(x > 20 \text{ cars}|\lambda t = 30) = \mathbf{0.9647}$$

Exponential distribution $X \sim \exp(\lambda)$

- **Definition**: The **exponential distribution** is often used to model time-to-failure of unrepairable components and systems because of its mathematical simplicity. The pdf for the exponential distribution is:

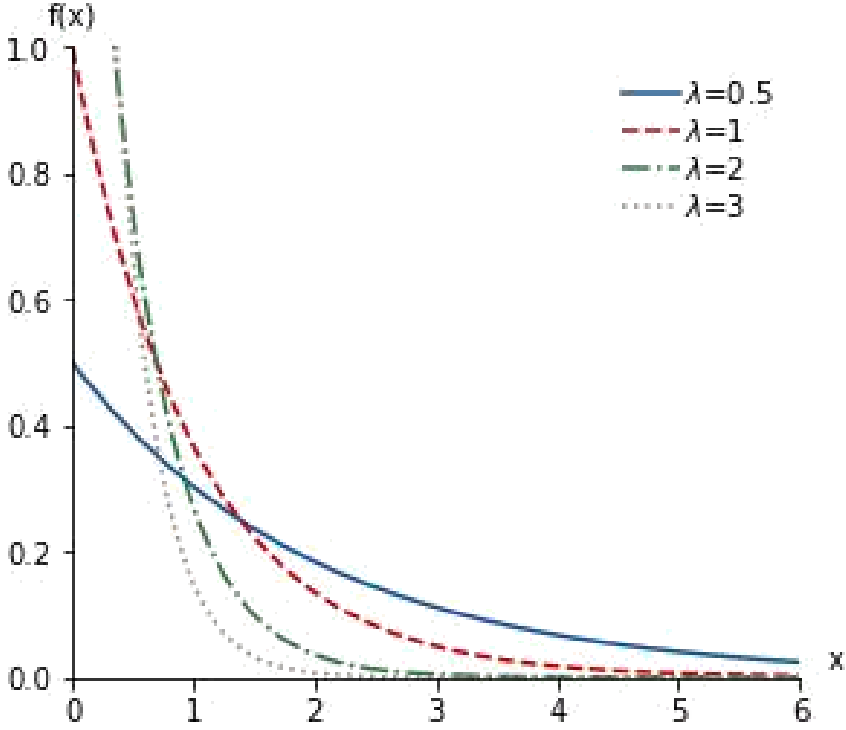
$$f(x|\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

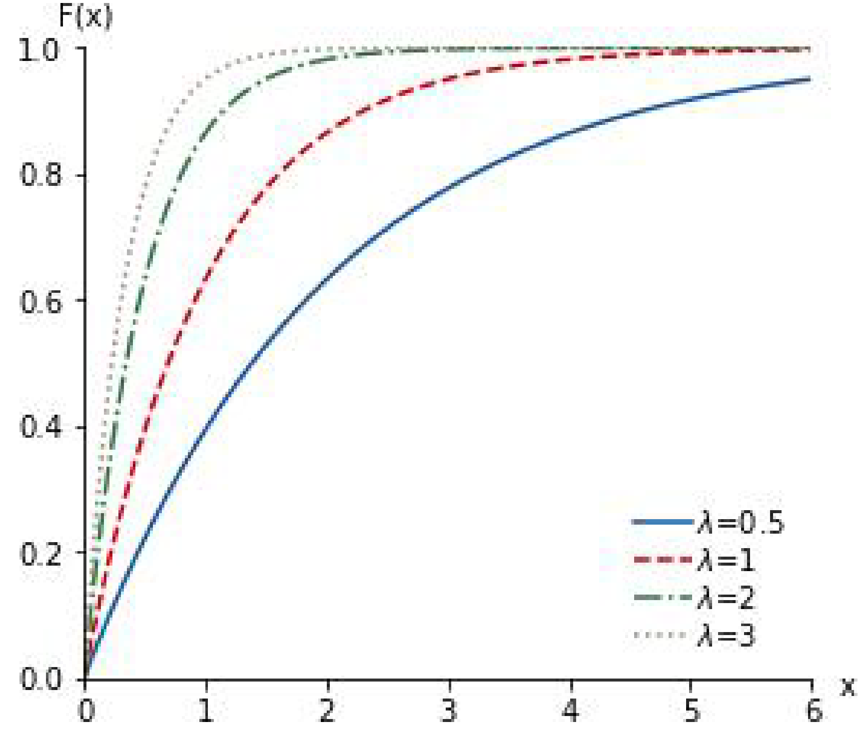
  Where:
  - λ – Rate (1/scale) parameter ($\lambda > 0$);

- The cdf for the exponential distribution is:

$$F(x|\lambda) = \begin{cases} 0 & x \leq 0 \\ \int_0^x \lambda e^{-\lambda x} dx = 1 - e^{-\lambda x} & x > 0 \end{cases}$$

Exponential distribution X~exp(λ)

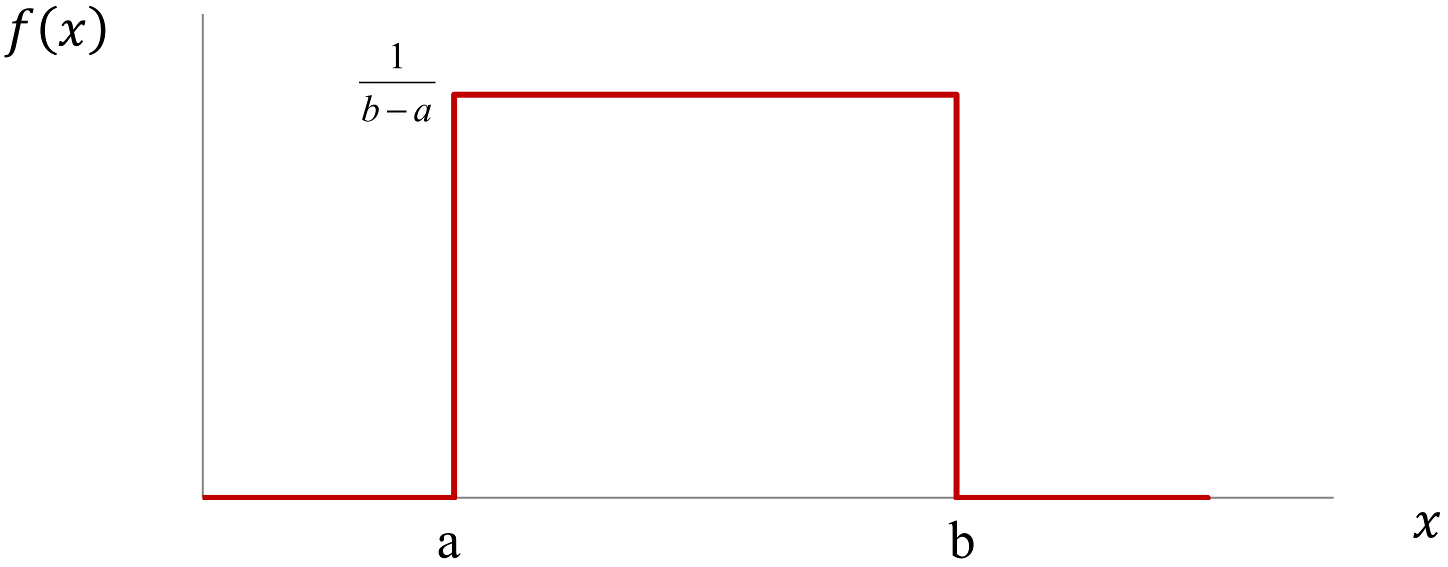




Uniform distribution *X~unif(a,b)*

- **Definition**: There is a continuous form of the **uniform distribution** sometimes called the **rectangular distribution**.

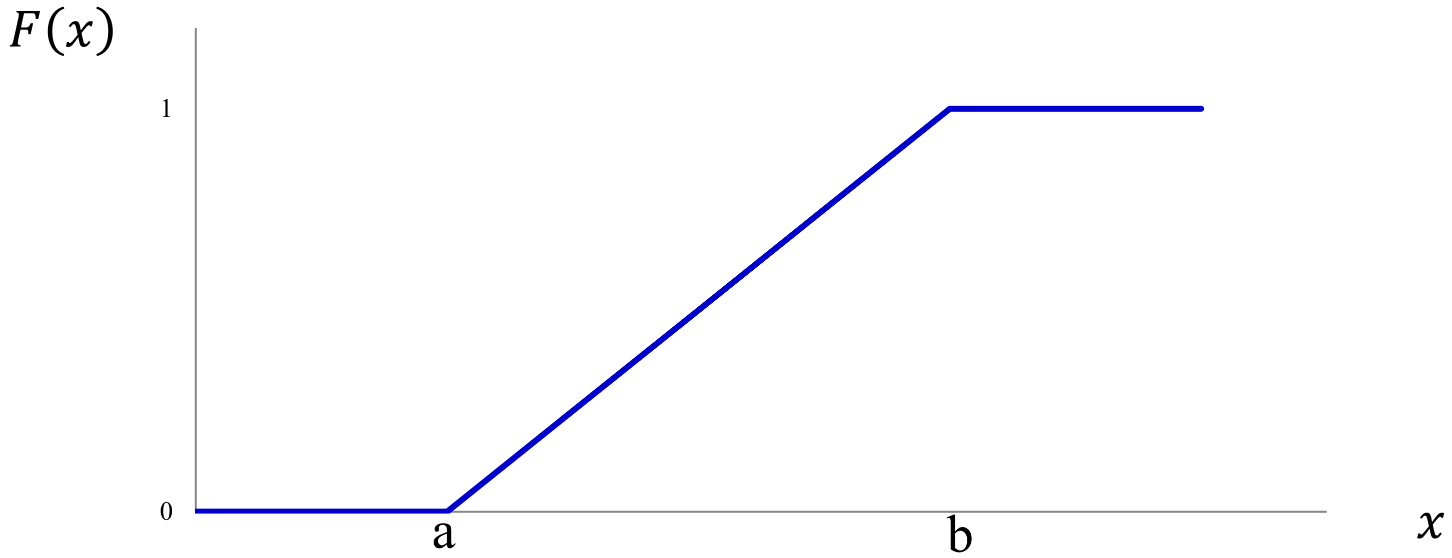
$$f(x|a,b) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$



Continuous uniform distribution

- **Definition**: The **cdf of the continuous uniform distribution** is

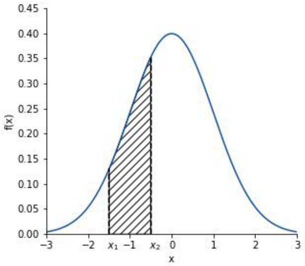
$$F(x|a,b) = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a \le x \le b \\ 1 & x > b \end{cases}$$



Normal distribution: $X \sim norm(\mu, \sigma)$

- The most well known and widely used pdf. The normal pdf sometimes called Gaussian distribution or Bell curve.
  - **Central Limit Theorem:** Means of non-normal variables are approximately normally distributed
  - The **pdf for a normal distribution** is:

$$f(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty$$



$$= \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)$$

Where:

- μ is the mean (location parameter) $-\infty < \mu < \infty$
- σ is the standard deviation (scale parameter) and $\sigma > 0$
- (and σ^2 is the variance)
- In Excel: = norm.dist(x, μ, σ, false)

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{\left[-\frac{1}{2}z^2\right]}$$

$\phi(z)$ is the standard normal distribution, where:

$$\boldsymbol{z} = \frac{x-\mu}{\sigma}$$

Normal Distribution

- The **cdf for a normal distribution** is:

$$F(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} \mathrm{e}^{\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right]}dt$$

$$= \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right) = \Phi(z)$$

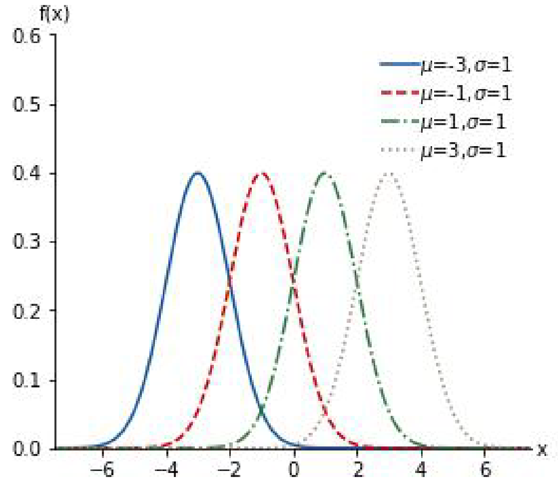
Where Φ is the standard normal CDF with $\mu = 0$ and $\sigma = 1$:

$$\Pr(Z \leq z) = \Phi(z)$$

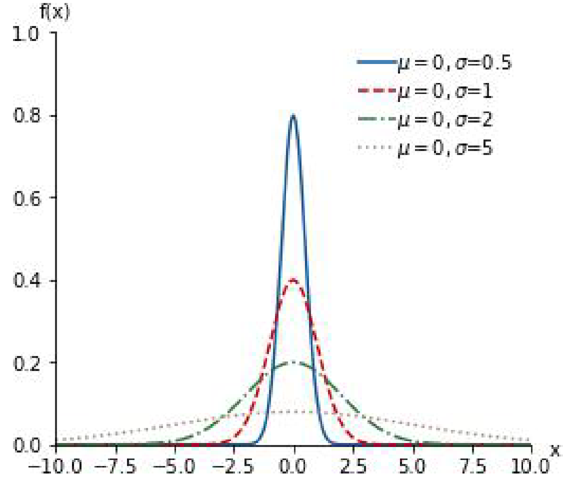
- All normal pdfs and cdfs can be transformed to standard form using the z transform: $Z = \frac{X-\mu}{\sigma} \rightarrow z = \frac{x-\mu}{\sigma}$
- **Appendix A** of the book gives a standard normal CDF table.

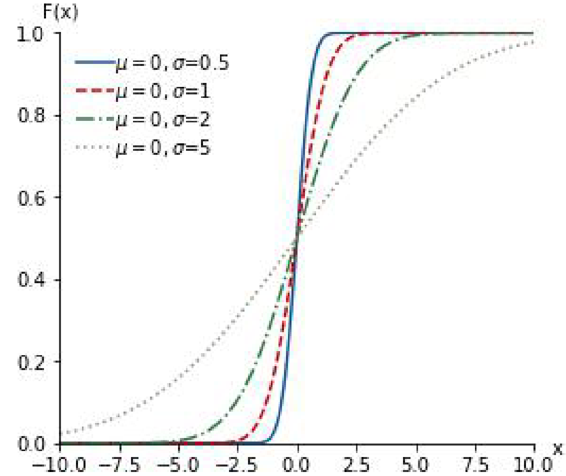
Normal distribution pdf, cdf

Varying μ and constant σ.



Varying σ and constant μ.





Normal distribution example

Example: Given a normal pdf with $\mu = 50$ and $\sigma = 10$. Find the probability that X is between 48 and 62, $\Pr(48 < x < 62)$.

Normal distribution example (cont.)

Example: Given a normal pdf with μ = 50 and σ = 10. Find the probability that *X* is between 48 and 62, $\Pr(48 < x < 62)$.

$$z_1 = \frac{48 - 50}{10} = -0.2 \quad z_2 = \frac{62 - 50}{10} = 1.2$$

$$\Pr(48 < x < 62) = \Pr(-0.2 < z < 1.2)$$

$$\Pr(z = 1.2) - \Pr(z = -0.2) = 0.885 - 0.421 = 0.464$$

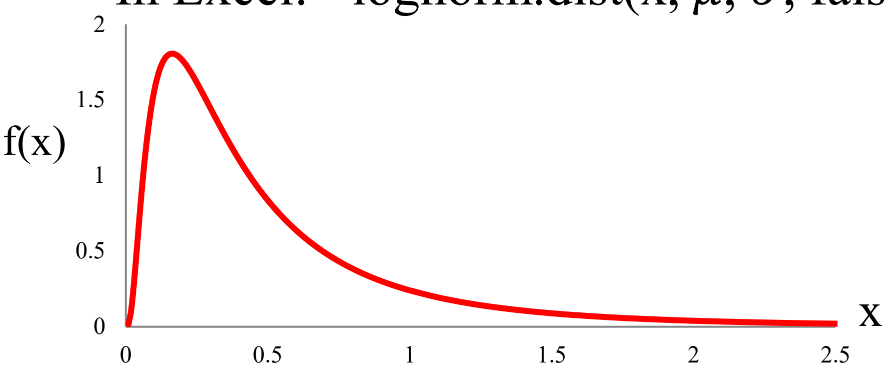
After class: solve using software and by hand

Lognormal distribution $X \sim lognorm(\mu, \sigma)$

- Definition: When the r.v. is defined as always positive it is said to be lognormally distributed if its ln is normally distributed.
  - That is: if x~lognorm, then ln(x) ~ norm.
- In reliability engineering it is used to represent cycles-or-time-to failure in fatigue. Generally used for product of a large number of i.i.d variables.

$$f(x|\mu,\sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{\left[-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2\right]} = \frac{1}{x\sigma}\phi\left(\frac{\ln x - \mu}{\sigma}\right) \qquad x > 0$$

  - μ is the scale parameter ($-\infty < \mu < \infty$); calculated as E(ln(x)))
  - σ is the shape parameter ($\sigma > 0$); calculated as stdev(ln(x))
- In Excel: =lognorm.dist(x, μ, σ, false); Can also use z-transform



$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{\left[-\frac{1}{2}z^2\right]}$$

$$\boldsymbol{z} = \frac{\ln(x) - \mu}{\sigma}$$

Lognormal distribution (cont.)

- Thus, the cdf for a lognormal distribution is,

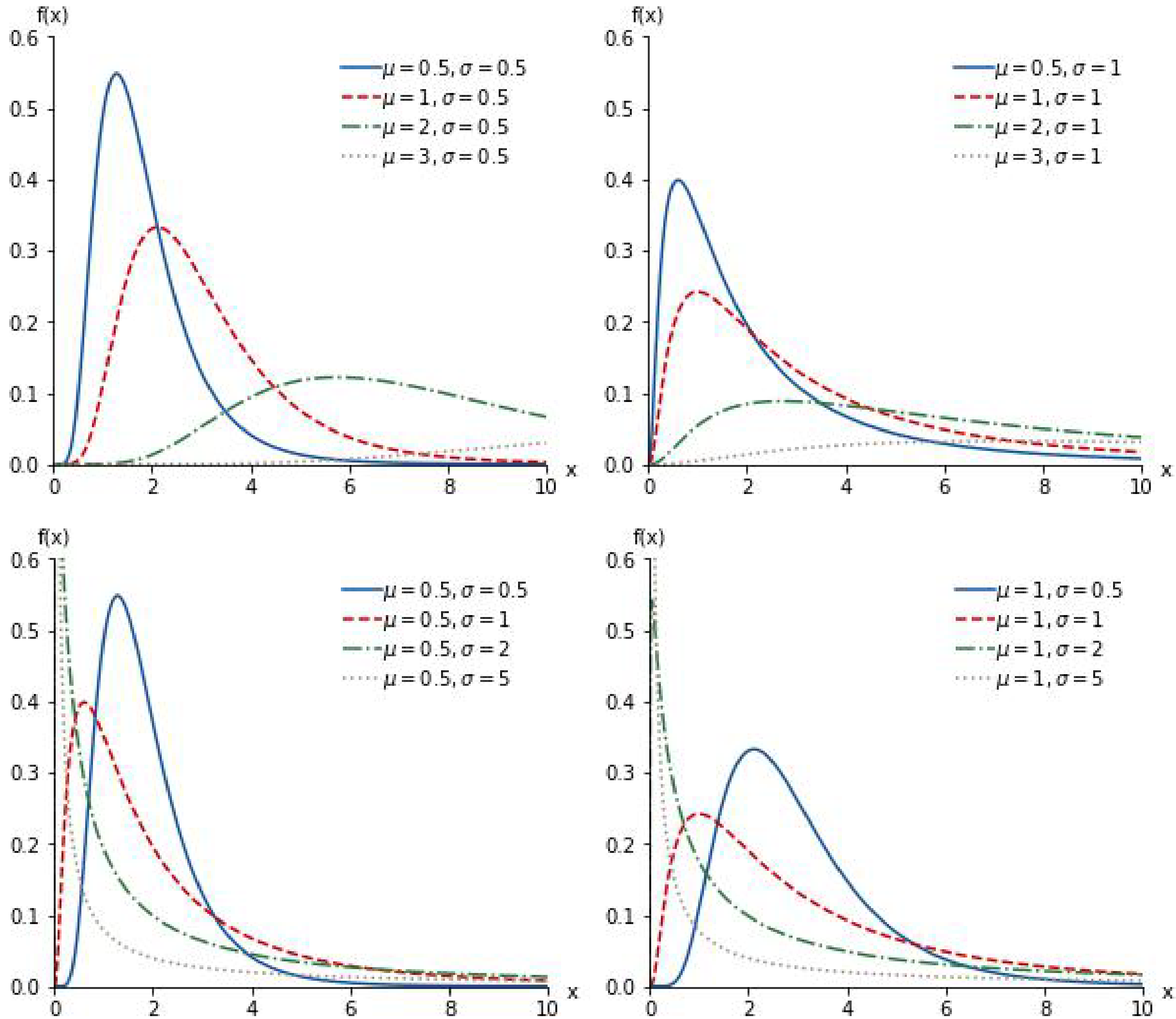
$$F(x|\mu,\sigma) = \int_0^x f(t)dt = \int_0^x \frac{1}{t\sigma\sqrt{2\pi}} e^{\left[-\frac{1}{2}\left(\frac{lnt-\mu}{\sigma}\right)^2\right]}dt$$

$$F(x) = \Phi\left(\frac{\ln(x) - \mu}{\sigma}\right)$$

Lognormal distribution pdf

Varying μ and constant σ.

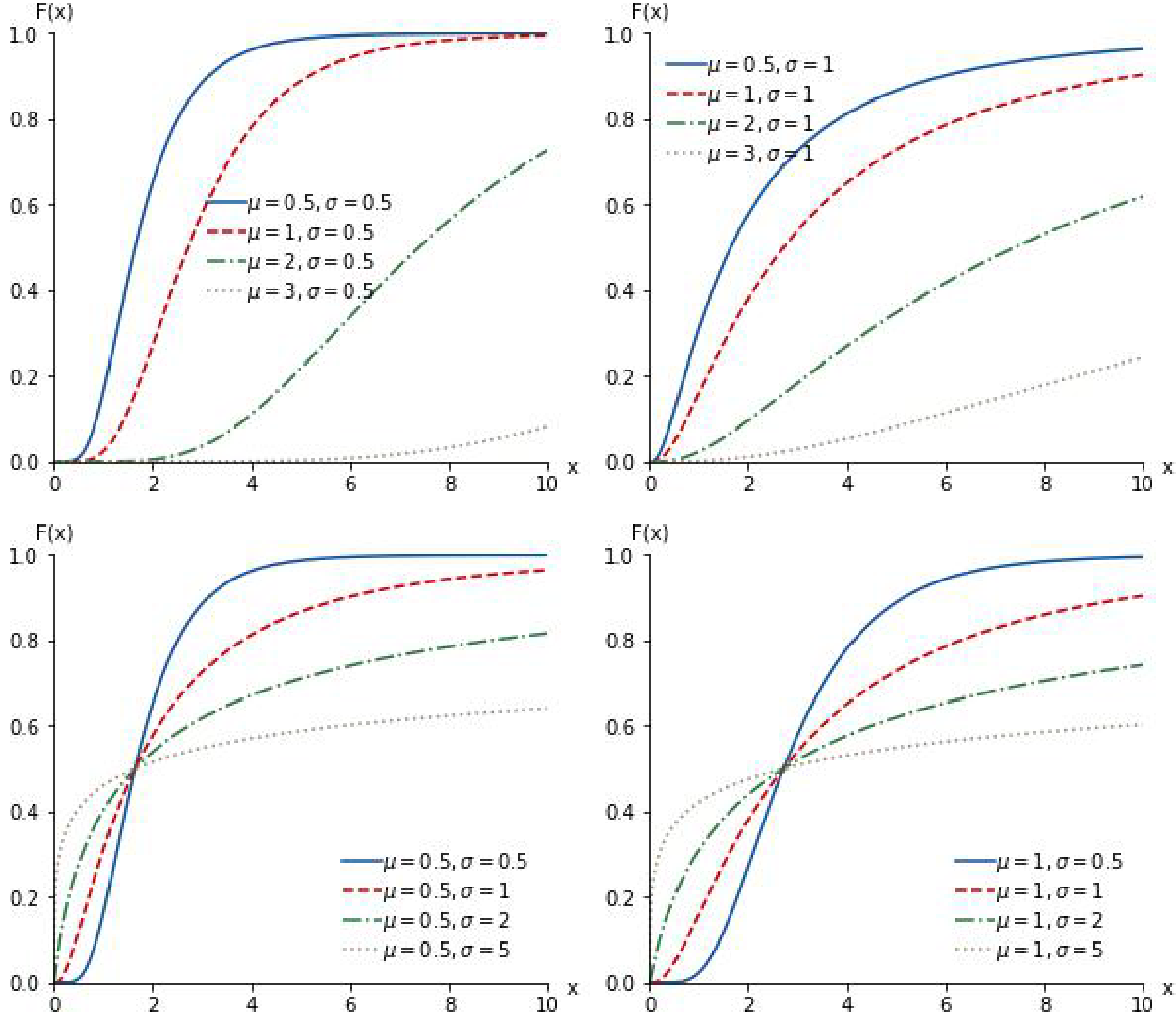
Varying σ and constant μ.



Lognormal distribution cdf

Varying μ and constant σ.

Varying σ and constant μ.



Lognormal distribution

Example: Given $\mu=1$ and $\sigma=1$, find the probability that lognormally distributed r.v. X will fall between 0.5 and 3.5.

Lognormal distribution

Example: Given μ=1, σ=1, find the probability that lognormally distributed *X* will fall between 0.5 and 3.5.

In Excel = lognorm.dist(3.5, 1, 1, true) – lognorm.dist (0.5, 1, 1, true)

$$z_1 = \frac{ln0.5 - 1}{1} = -1.69 \qquad z_2 = \frac{ln3.5 - 1}{1} = 0.25$$

By hand, using Appendix A tables:

$$\Pr(0.5 < X < 3.5) = \Pr(-1.69 < Z < 0.25)$$

$$= \Phi(0.25) - \Phi(-1.69) = 0.59871 - 0.04551 = \mathbf{0.553}$$

(Using software without rounding: **=0.5545**)

Weibull distribution $X{\sim}weibull(\alpha, \beta)$

- **Definition**: The **Weibull distribution** is a very flexible and popular distribution that can be used in a variety of situations. Primarily for failure or life-based reliability problems.

$$f(x|\alpha, \beta) = \frac{\beta x^{\beta-1}}{\alpha^{\beta}} \mathrm{e}^{\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right]} \quad x \geq 0$$

where,

- α is the scale parameter ($\alpha > 0$)
- β is the shape parameter ($\beta > 0$)

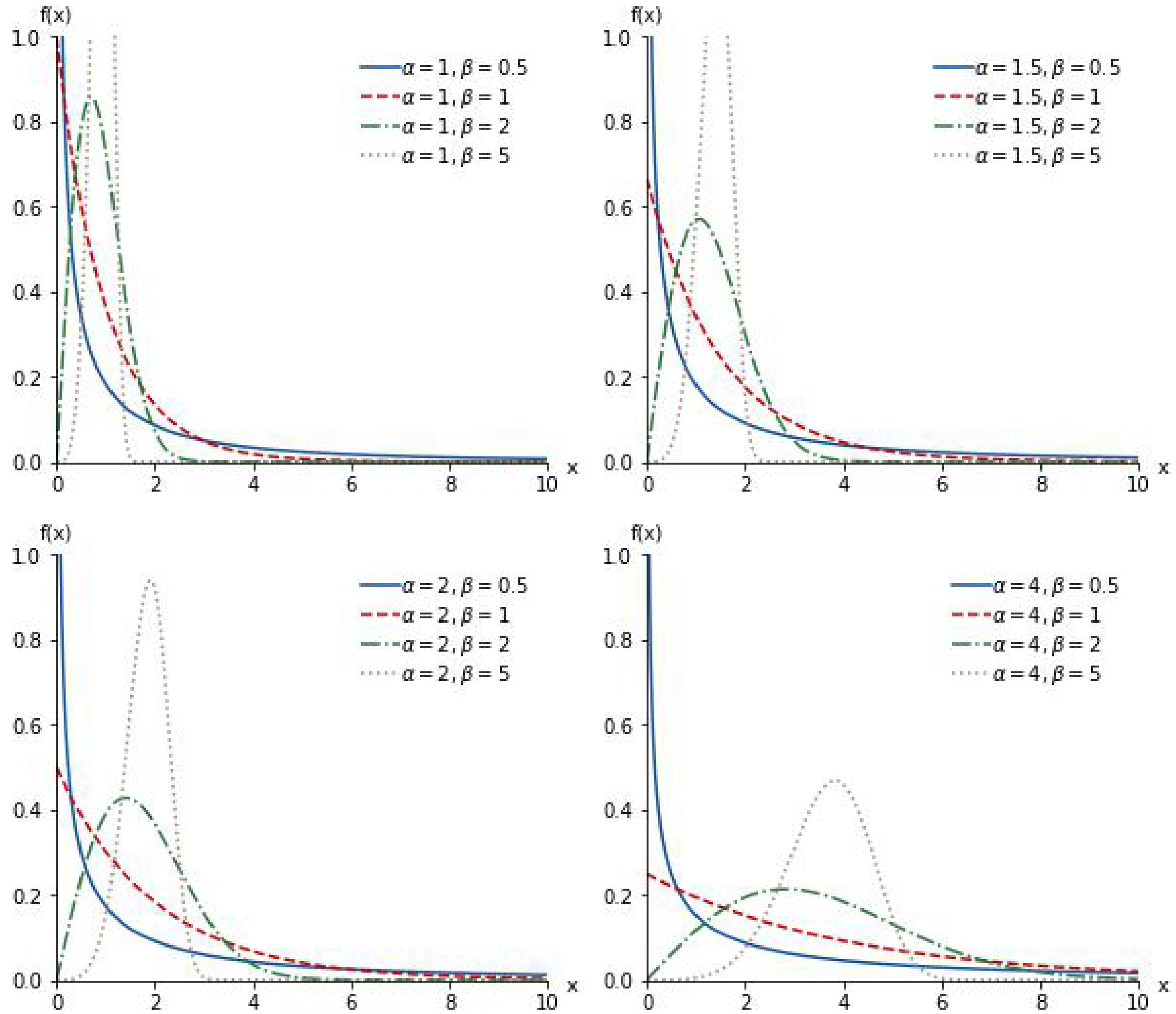
Weibull distribution

Example: For a Weibull distribution with the scale parameter α and shape parameter β, the cdf is

$$F(x|\alpha,\beta) = \begin{cases} 1 - e^{\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right]}, & x, \alpha, \beta > 0 \\ 0 & \text{otherwise} \end{cases}$$

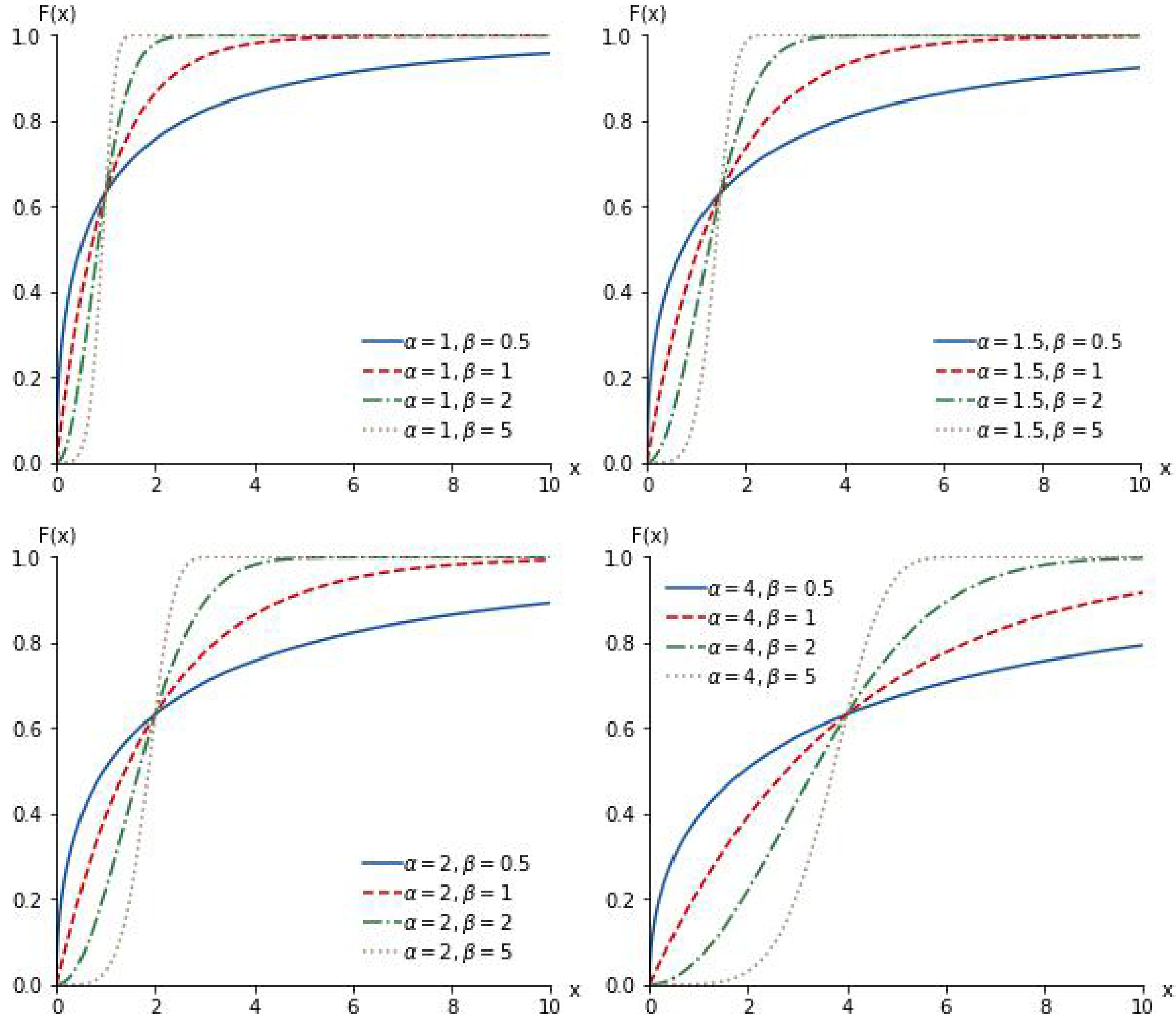
Weibull distribution pdf

Each frame has constant α and varying β.



Weibull distribution cdf

Each frame has constant α and varying β.



Weibull distribution

Example: When the shape parameter $\beta = 1$, determine what other pdf this distribution becomes and what is its governing parameter value?

$$f(x|\alpha, 1) = \frac{1x^{1-1}}{\alpha^1} \mathrm{e}^{\left[-\left(\frac{x}{\alpha}\right)^1\right]} = \frac{1}{\alpha} \mathrm{e}^{-\frac{x}{\alpha}}$$

- What does this remind you of? Compare this to

$$f(x) = \lambda e^{(-\lambda x)}$$

- So, the Weibull pdf with $\beta = 1$ becomes an exponential distribution with the parameter $\lambda = \frac{1}{\alpha}$.

Gamma distribution $X \sim gamma(\alpha, \beta)$

- **Definition**: The **gamma distribution** is a generalized form of the exponential distribution. It represents the sum of α independent exponential variables. However, it bears a closer similarity to the Poisson distribution. Two forms depending on the shape parameter:

$$f(x|\alpha,\beta) = \begin{cases} \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{\left(-\frac{x}{\beta}\right)} & \alpha \text{ is continuous} \\ \dfrac{1}{\beta^{\alpha}(\alpha-1)!} x^{\alpha-1} e^{\left(-\frac{x}{\beta}\right)} & \alpha \text{ is an integer} \end{cases}$$

where,

- α – shape parameter and $\alpha > 0$
- β – scale parameter and $\beta > 0$
- Range – $x \geq 0$

The mean and variance is given as: $\mathrm{E}(X) = \alpha\beta$ and $Var(X) = \alpha\beta^2$

In Excel: Gamma.dist(x, α, β, *false*)

The chi squared distribution is a gamma distribution with $\beta = 2,\ \alpha = df/2$ where df is the number of degrees of freedom.

Gamma function (used in gamma distribution)

- **Definition**: The $\Gamma(\alpha)$ term in the continuous form is known as the **gamma function** which is given as,

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{(-x)} dx$$

- Values of the function are listed in the following Table.
  - Alternately, in both Excel and Matlab enter gamma(α)

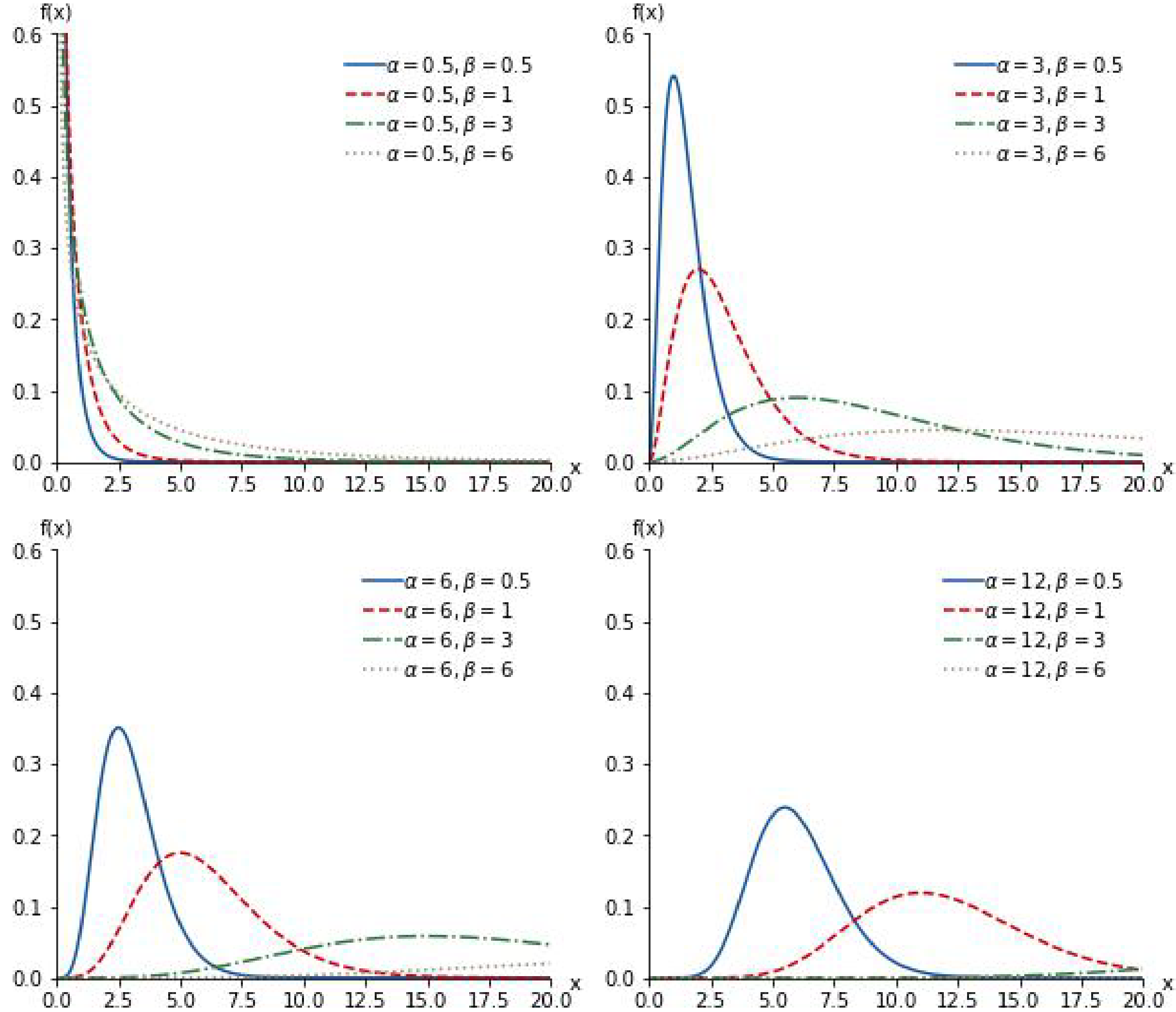
Gamma function table

α	$\Gamma(\alpha)$	α	$\Gamma(\alpha)$	α	$\Gamma(\alpha)$	α	$\Gamma(\alpha)$	α	$\Gamma(\alpha)$
1.0000	1.0000	1.4000	0.8873	1.8000	0.9314	3.0000	2.0000	5.0000	24.0000
1.0200	0.9888	1.4200	0.8864	1.8200	0.9368	3.1000	2.1976	5.1000	27.9318
1.0400	0.9784	1.4400	0.8858	1.8400	0.9426	3.2000	2.4240	5.2000	32.5781
1.0600	0.9687	1.4600	0.8856	1.8600	0.9487	3.3000	2.6834	5.3000	38.0780
1.0800	0.9597	1.4800	0.8857	1.8800	0.9551	3.4000	2.9812	5.4000	44.5988
1.1000	0.9514	1.5000	0.8862	1.9000	0.9618	3.5000	3.3234	5.5000	52.3428
1.1200	0.9436	1.5200	0.8870	1.9200	0.9688	3.6000	3.7170	5.6000	61.5539
1.1400	0.9364	1.5400	0.8882	1.9400	0.9761	3.7000	4.1707	5.7000	72.5276
1.1600	0.9298	1.5600	0.8896	1.9600	0.9837	3.8000	4.6942	5.8000	85.6217
1.1800	0.9237	1.5800	0.8914	1.9800	0.9917	3.9000	5.2993	5.9000	101.2702
1.2000	0.9182	1.6000	0.8935	2.000	1.000	4.0000	6.0000	6.0000	120.0000
1.2200	0.9131	1.6200	0.8959	2.1000	1.0465	4.1000	6.8126	6.1000	142.4519
1.2400	0.9085	1.6400	0.8986	2.2000	1.1018	4.2000	7.7567	6.2000	169.4061
1.2600	0.9044	1.6600	0.9017	2.3000	1.1667	4.3000	8.8553	6.3000	201.8133
1.2800	0.9007	1.6800	0.9050	2.4000	1.2422	4.4000	10.1361	6.4000	240.8338
1.3000	0.8975	1.7000	0.9086	2.5000	1.3293	4.5000	11.6317	6.5000	287.8853
1.3200	0.8946	1.7200	0.9126	2.6000	1.4296	4.6000	13.3813	6.6000	344.7019
1.3400	0.8922	1.7400	0.9168	2.7000	1.5447	4.7000	15.4314	6.7000	413.4075
1.3600	0.8902	1.7600	0.9214	2.8000	1.6765	4.8000	17.8379	6.8000	496.6061
1.3800	0.8885	1.7800	0.9262	2.9000	1.8274	4.9000	20.6674	6.9000	597.4941

For integer values of α, $\Gamma(\alpha+1) = \alpha!$
Note that $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$; Excel and Matlab: gamma(α)

Gamma distribution pdf

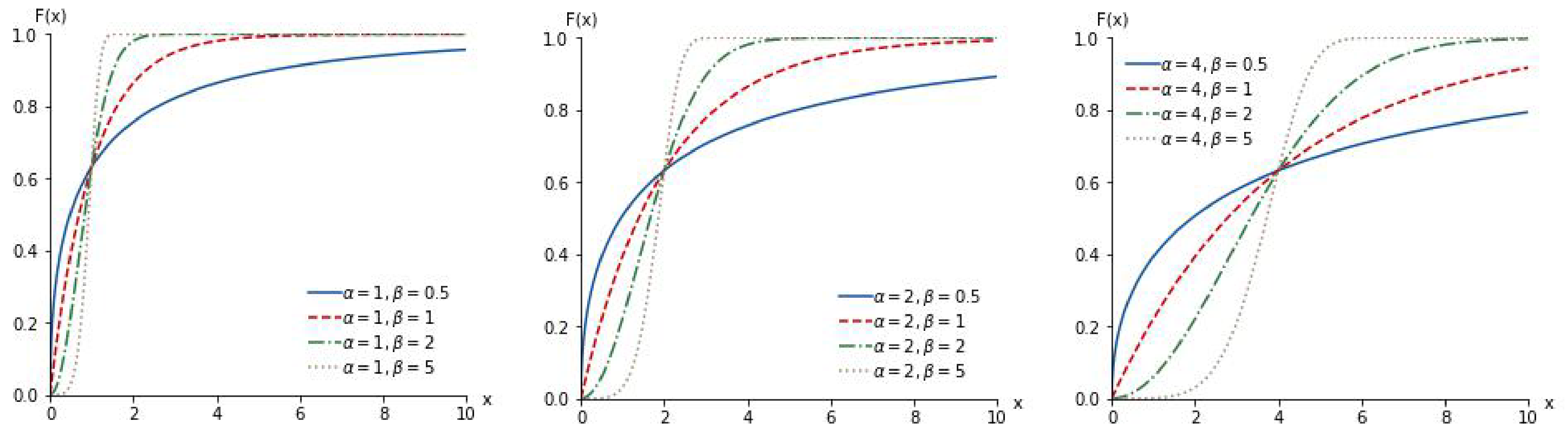
Each frame has constant α and varying β.



Gamma distribution cdf

- The cdf for the gamma distribution is

$$F(x|\alpha,\beta)=\begin{cases}\dfrac{1}{\beta^{\alpha}\Gamma(\alpha)}\displaystyle\int_0^x x^{\alpha-1}\,e^{\left(-\frac{x}{\beta}\right)}dx & x,\alpha,\beta>0 \text{ and is } \alpha\text{continuous}\\ 1-e^{\left(-\frac{x}{\beta}\right)}\displaystyle\sum_{n=0}^{\alpha-1}\frac{x^{\alpha-1}}{\beta^{n}n!} & x,\alpha,\beta>0 \text{ and is } \alpha \text{ an integer}\\ 0 & \text{otherwise}\end{cases}$$



Note: The Chi-Sq. distribution is a special case of the Gamma distribution

Beta distribution $X \sim beta(\alpha, \beta)$

- **Definition**: Use the **beta distribution** for r.v.s that are distributed over the finite interval between 0 and 1.

$$f(x|\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$$

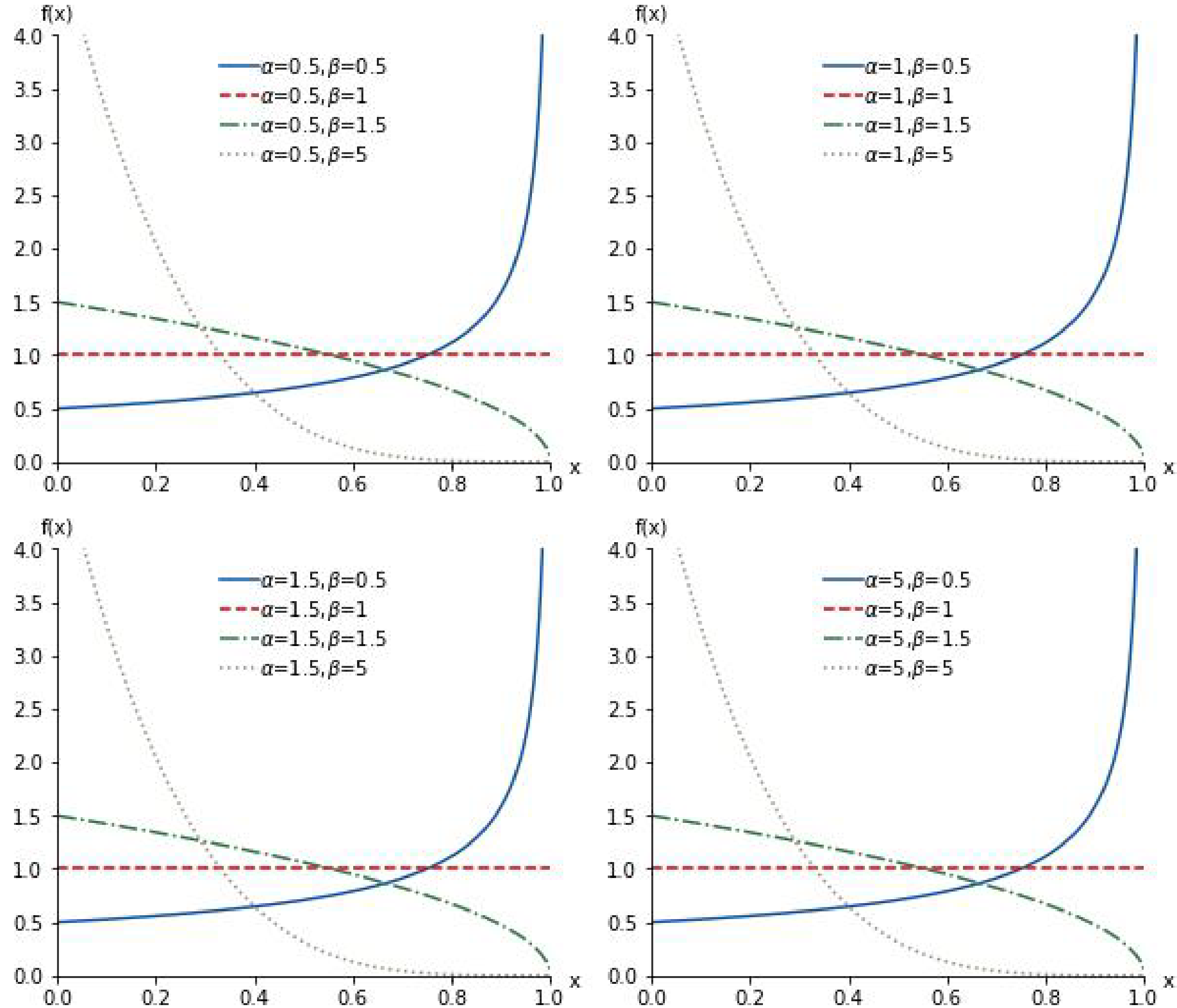
  Where:
  - α – shape parameter and $\alpha > 0$
  - β – shape parameter and $\beta > 0$
  - Range – $0 \leq x \leq 1$
  - Excel: $Beta.Dist(X, \alpha, \beta, false)$

- Like the gamma distribution, it too uses the gamma function. The mean and variance of the beta distribution

$$E(X) = \frac{\alpha}{\alpha+\beta} \text{ and } \mathrm{Var}(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Beta distribution pdf

Each frame has constant α and varying β.



Beta distribution (cont.)

- The cdf of the beta distribution is

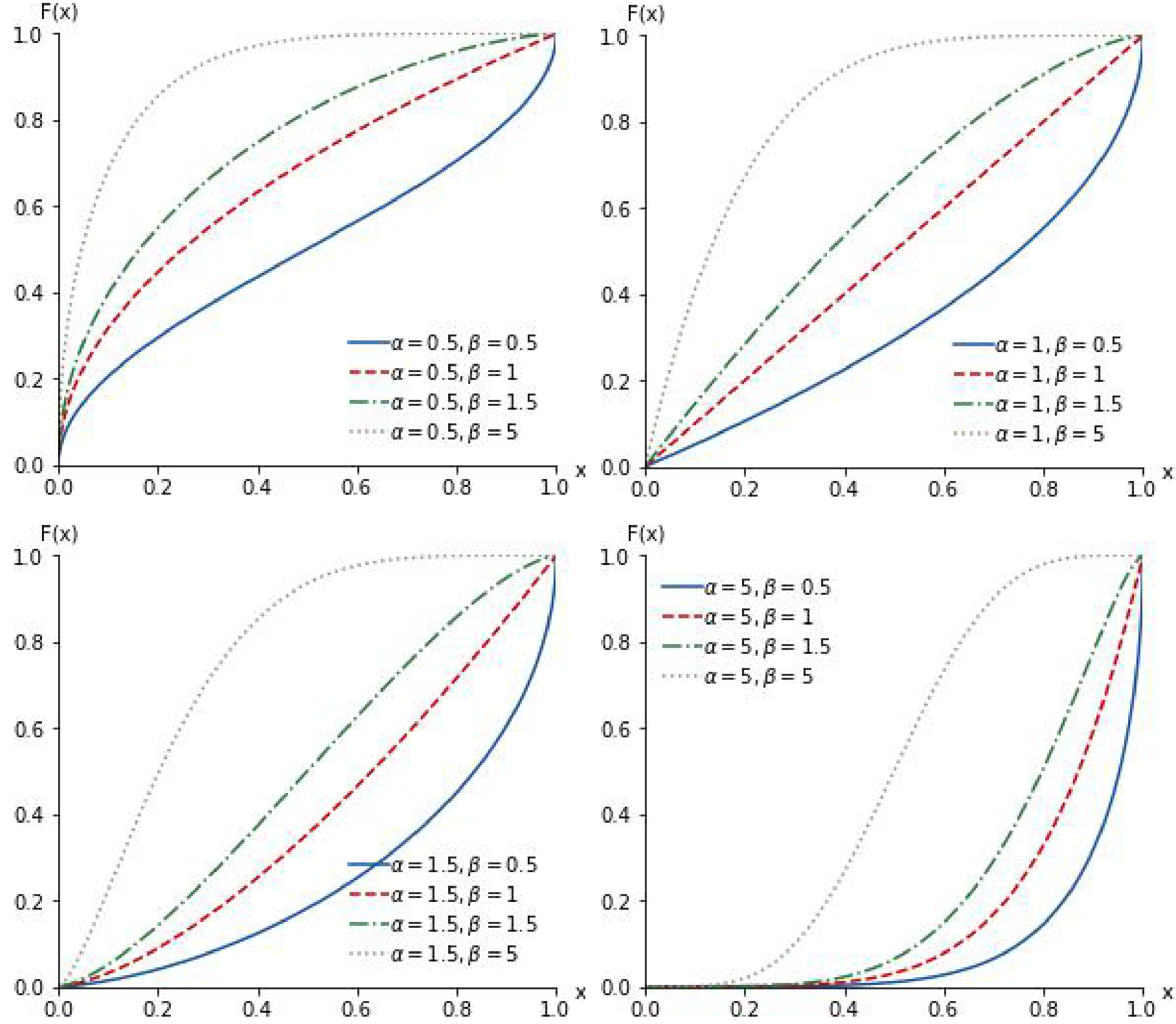
$$F(x|\alpha,\beta) = \begin{cases} 0 & x < 0 \\ \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \displaystyle\int_0^x x^{\alpha-1}(1-x)^{\beta-1}dx & 0 \le x \le 1 \\ 1 & x > 1 \end{cases}$$

Where:

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{(-x)} dx$$

Beta distribution cdf

Each frame has constant α and varying β.



Truncated distributions

- **Truncation** arises when the r.v. cannot exist in some range of a distribution.
- A **truncated distribution** is the conditional distribution that results from restricting the domain of another probability distribution.
- The following **general formulas apply to truncated distribution functions**, where $f_0(x)$ and $F_0(x)$ are the pdf and cdf of the non-truncated distribution.

PDF:

$$f(x) = \begin{cases} \dfrac{f_0(x)}{F_0(b) - F_0(a)} & for\ x \in (a, b] \\ 0 & otherwise \end{cases}$$

CDF:

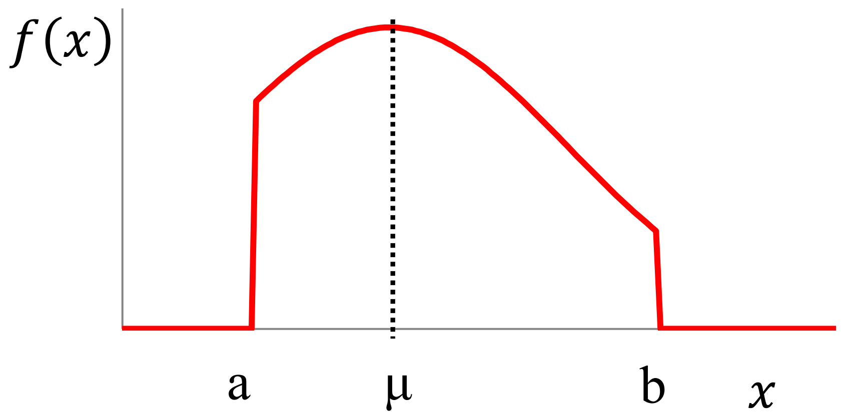
$$F(x) = \begin{cases} 0 & x \le a \\ \dfrac{\int_a^x f_0(t)dt}{F_0(b) - F_0(a)} & x \in (a, b] \\ 1 & x > b \end{cases}$$

Truncated normal distribution

- $a \le x \le b$
- Pdf:

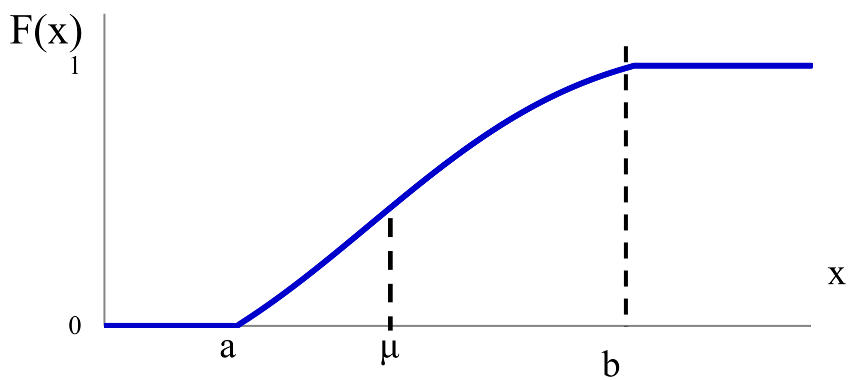
$$f(x) = \frac{f_{norm}(x)}{F_{norm}(x = b_U) - F_{norm}(x = a_L)}$$

$$= \frac{1}{\sigma} \frac{\phi(z_x)}{\Phi(z_b) - \Phi(z_a)}$$



- Cdf:

$$F(x) \begin{cases} 0 & x < a \\ \dfrac{F_{Norm}(x) - F_{Norm}(x = a)}{F_{Norm}(x = b) - F_{Norm}(x = a)} = \dfrac{\Phi(z_x) - \Phi(z_a)}{\Phi(z_b) - \Phi(z_a)} & a \le x \le b \\ 1 & x > b \end{cases}$$



Multivariate distributions

- **Definition**: Let's say that several random variables take the following values,

$$(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) \text{ or } (x_1, x_2, \dots, x_n)$$

- Then a **joint (multivariate) pdf** is a function $f(x_1, x_2, \dots, x_n)$ that satisfies the following conditions:

  i. $f(x_1, x_2, \dots, x_n) \geq 0$ and $-\infty < x_i < \infty$ for $i = 1,2, \dots, n$

  ii. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n = 1$

- **Definition**: The **marginal pdf** of X_i is defined by,

$$f_i(x_i) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_{i-1} dx_{i+1} \dots dx_n$$

Multivariate distributions

- **Definition**: The **conditional pdf** of X_1 given X_2 is denoted by $g(x_1|x_2)$ and defined as,

$$g(x_1|x_2) = \frac{f(x_1, x_2)}{f_2(x_2)}, f_2(x_2) \neq 0$$

- Two random variables are **independent** if their pdf follows,

$$f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2)$$

- Similarly,

$$f(x_1, x_2, \dots, x_n) = f_1(x_1) \cdot f_2(x_2) \cdot \dots \cdot f_n(x_n)$$