

# 可靠度資料分析

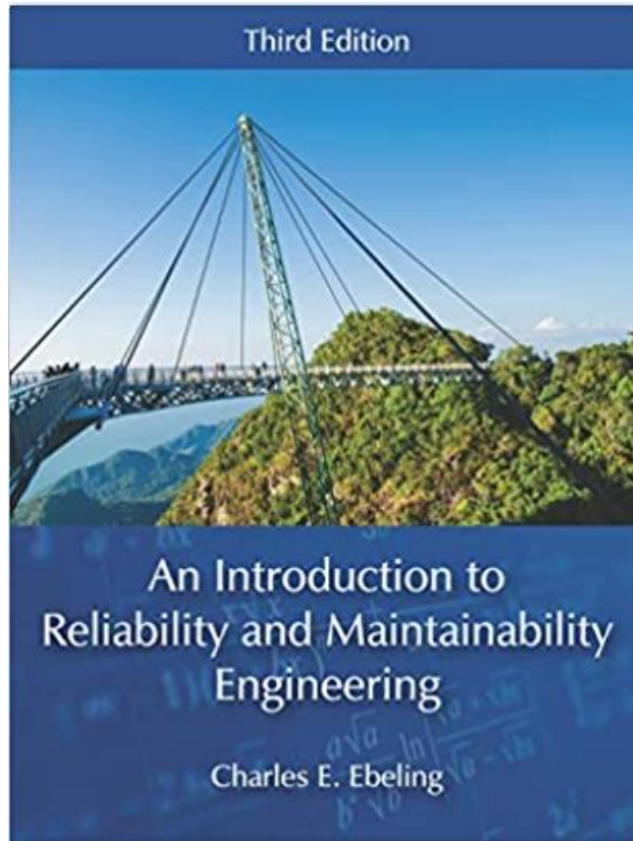
# Reliability Data Analysis

許舒涵 (Shu-han Hsu)

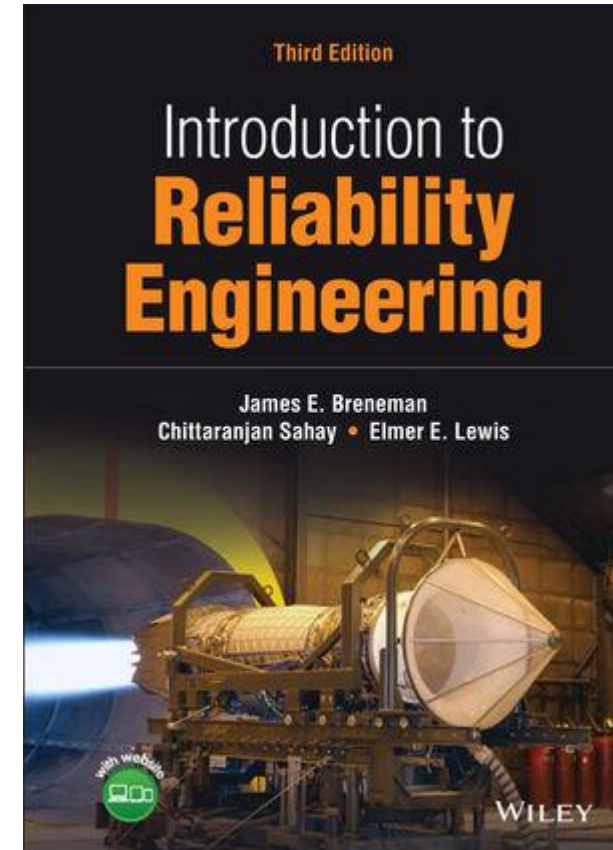
成功大學 資訊工程系

**Lecture 2 – Probability & Statistics Concepts**

# Textbook



Can only buy e-book for \$55US (~\$1674 NT)



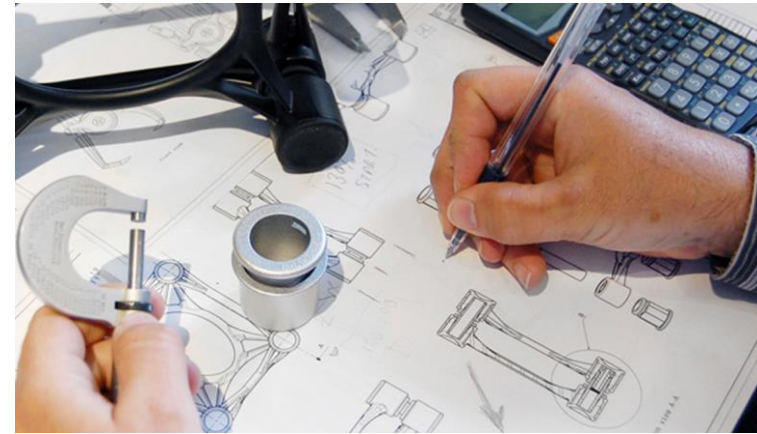
華通書局  
\$1793 NT for print version

# Recap: What is Reliability?

- **Reliability** is the **probability** that a product will operate or a service will be provided properly for a specified period of time (known as the design life) under the intended operating conditions (designated temperature, load, speed, etc.) without failure.
- **Reliability Engineering** attempts to study, characterize, measure, and analyze system failures in order to improve their operational use by increasing their **design life** and reducing the likelihood of unexpected failures, and **downtime**, thereby increasing **availability**.

# How reliability can be used

- Designing a device / product / system
- Correcting product reliability/warranty issue

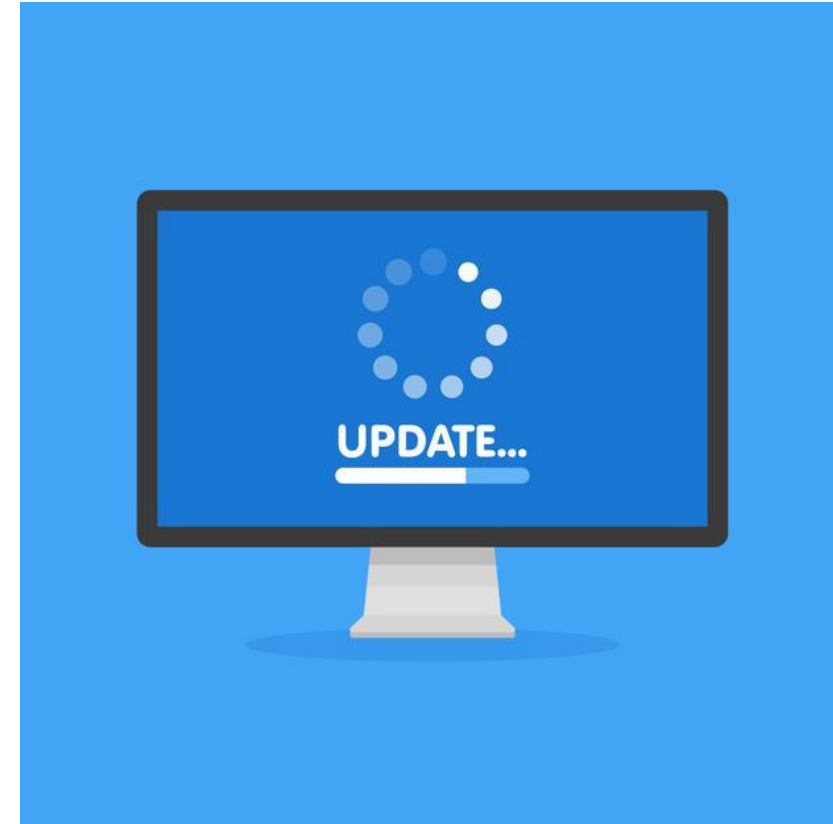


<https://www.sparkinnovations.com/product-design/>



# Designing a Device/Product

- Incremental Update
- Significant Update
- New Technology / Material



# Incremental Update

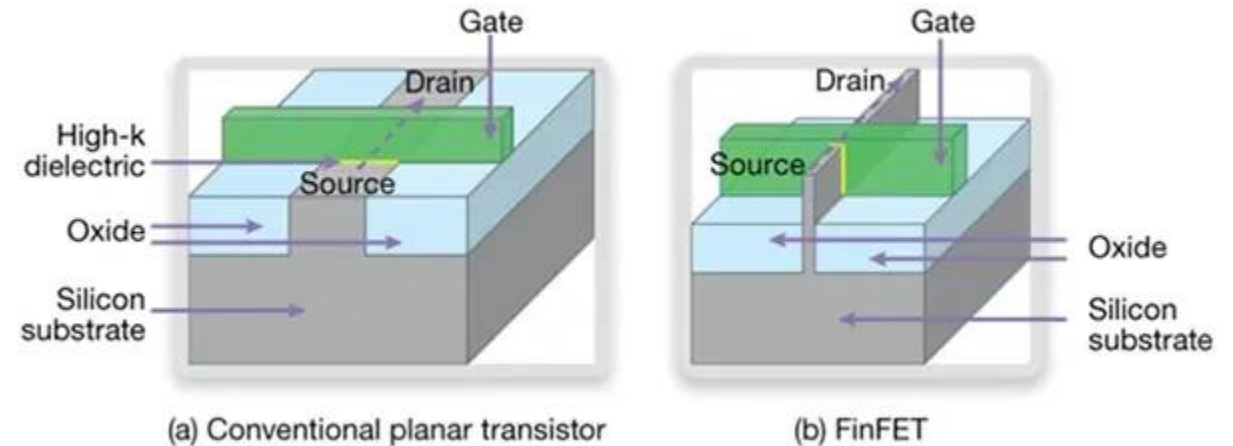
- Typical Business Rhythm
- Well Defined Schedule / Low Risk
- Design Guidelines
- Routine Reliability Testing
- Examples
  - 45nm -> 32 nm process node
  - Liquid Crystal Display (LCD) 720 ->1080
  - iPhone updates



# Significant Product Change

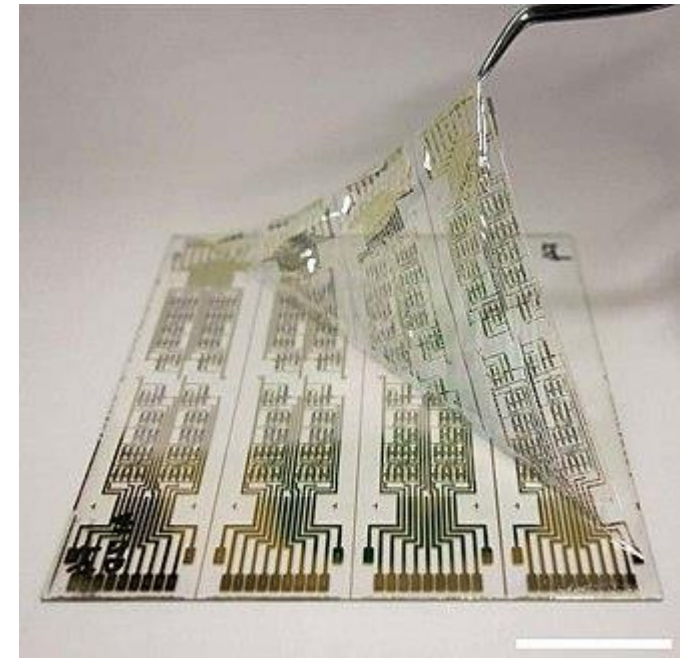
- Business Decision to Disrupt Market
- Multiple Design Changes
- Risks: Cost, Schedule, Performance
- Design Guidelines Adjusted
- Highly Accelerated Life Testing
- Examples
  - FinFET process
  - Introduction of iPhone

## Semiconductor Process Change



# New Technology / Material

- Investment Driven by:
  - Physical limitations of Existing Technology
  - Entry of Company in New Market
- Reliability / Failure Mechanisms Unknown
- Risks
  - Warranty costs
  - Company reputation
- Examples
  - Organic Transistor
  - Organic Light Emitting Diode (OLED Display)



Organic CMOS logic circuit

[https://en.wikipedia.org/wiki/Organic\\_field-effect\\_transistor](https://en.wikipedia.org/wiki/Organic_field-effect_transistor)

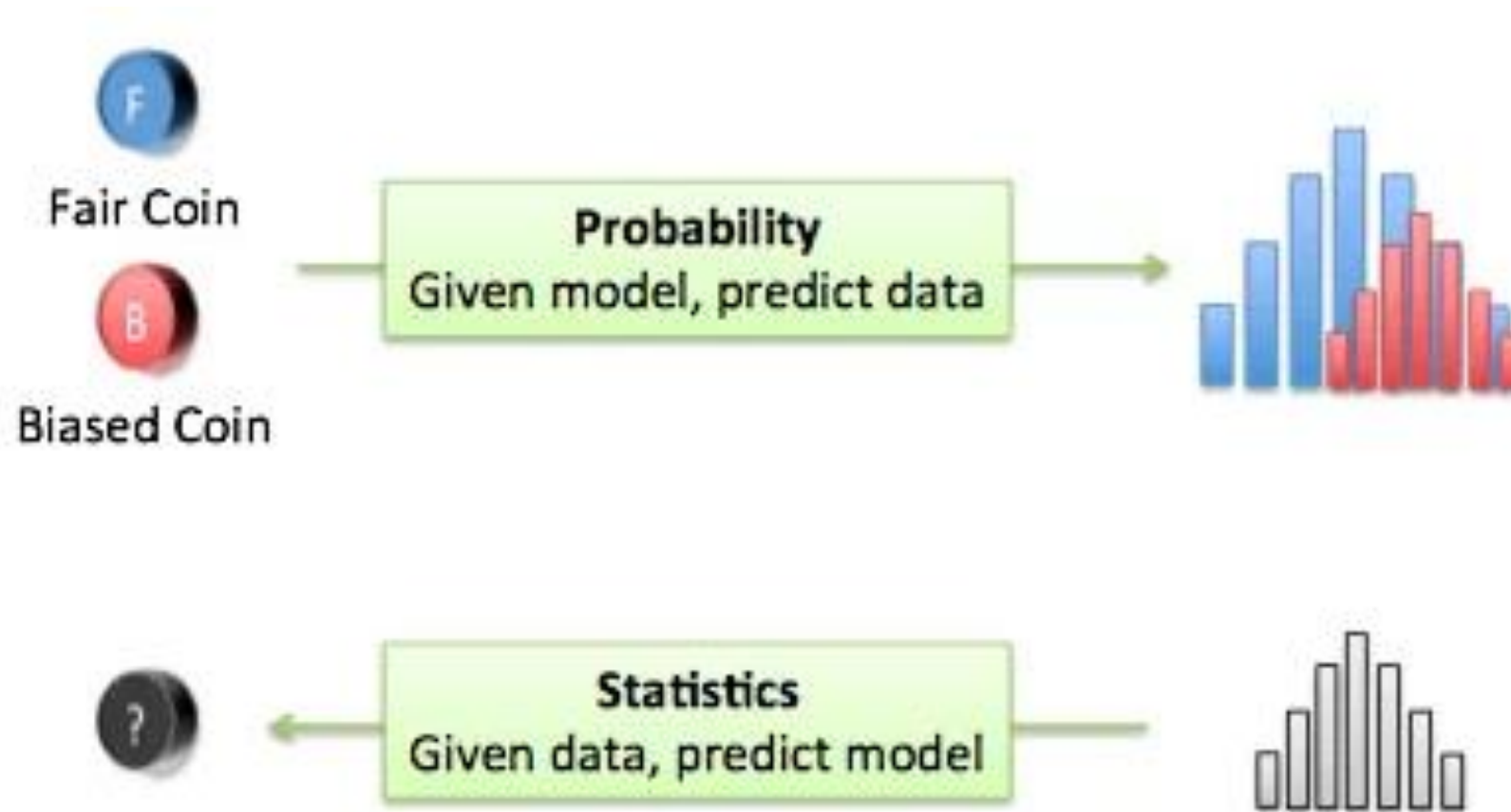


# Product Field Reliability Issue

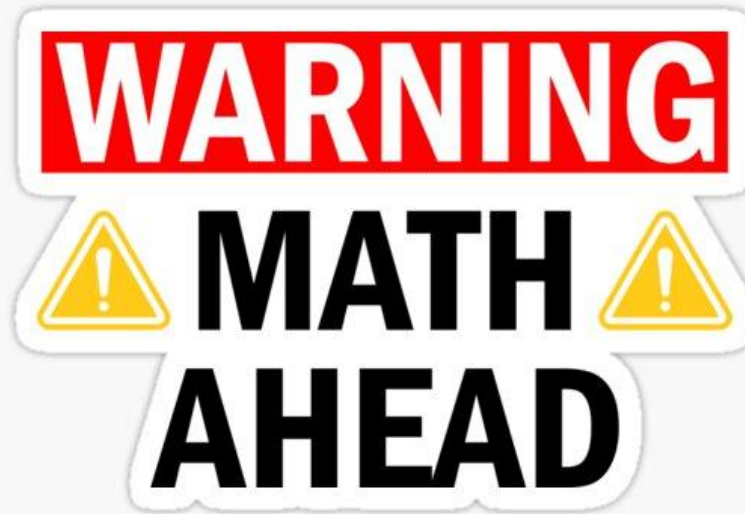
- Determine Root Cause
  - Failure analysis
  - Unbiased perspective
- Identify Corrective Action
  - Test solution thoroughly
  - Balance solution with severity of situation
- Preventative Action
  - Often overlooked
  - Improve design process



# Why Probability & Statistics?

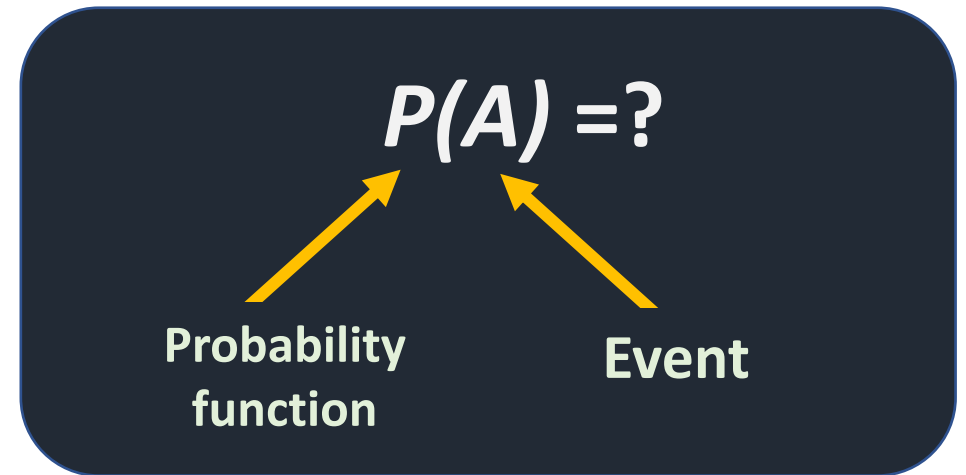


# Probability Concepts Review



# Basic Probability Model

- A probability model consists of an experiment which produces exactly one out of several mutually exclusive (doesn't occur at same time) outcomes
- Essential elements are:
  - Sample space  $S$  (some texts use  $\Omega$  or  $U$ )
    - Collection or list (set) of possible outcomes
  - Probability law  $P(\cdot)$ 
    - assigns a “likelihood” to different events
      - Event: a subset of the sample space
    - Ex. Probability of  $A$  is given by  $P(A)$



# Example: Roll of a dice

- Consider a fair six-sided die. The experiment is rolling the dice. The sample space is given by

$$\Omega = \{ \begin{array}{|c|} \hline \bullet \\ \hline \end{array}, \begin{array}{|c|} \hline \bullet \\ \hline \bullet \\ \hline \end{array}, \begin{array}{|c|} \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \end{array}, \begin{array}{|c|} \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \end{array}, \begin{array}{|c|} \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \end{array}, \begin{array}{|c|} \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \end{array} \}$$

- Possible events might include
  - the result is a “1”:  $\{ \begin{array}{|c|} \hline \bullet \\ \hline \end{array} \}$
  - the result is odd:  $\{ \begin{array}{|c|} \hline \bullet \\ \hline \end{array}, \begin{array}{|c|} \hline \bullet \\ \hline \bullet \\ \hline \end{array}, \begin{array}{|c|} \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \end{array} \}$
  - the result is even:  $\{ \begin{array}{|c|} \hline \bullet \\ \hline \bullet \\ \hline \end{array}, \begin{array}{|c|} \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \end{array}, \begin{array}{|c|} \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \end{array} \}$
  - the result is less than or equal to “3”:  $\{ \begin{array}{|c|} \hline \bullet \\ \hline \end{array}, \begin{array}{|c|} \hline \bullet \\ \hline \bullet \\ \hline \end{array}, \begin{array}{|c|} \hline \bullet \\ \hline \bullet \\ \hline \bullet \\ \hline \end{array} \}$
- There are  $2^6$  different possible events if we allow  $\Omega$  and  $\emptyset$  to count as events

# Example: Roll of a dice

- Since the die is “fair”, a natural probability law is to assign each of the six possible outcomes the same value

$$P(\{\square\bullet\}) = P(\{\bullet\square\}) = \dots P(\{\square\square\square\square\square\square\}) = \frac{1}{6}$$

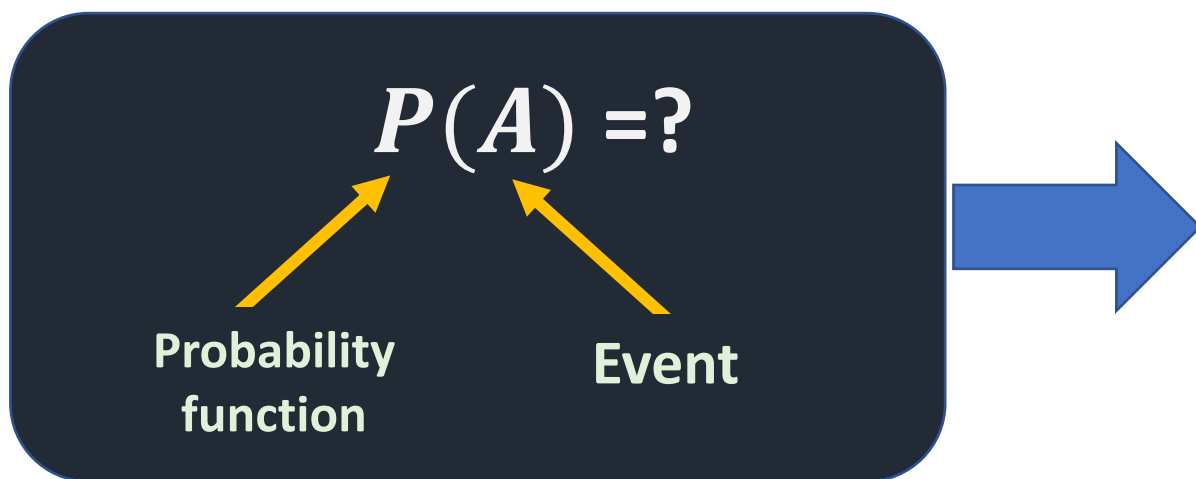
- It is then straightforward to compute the corresponding probability of different events:

$$\bullet P(\{\square\bullet, \bullet\square\square, \square\square\square\square\}) = \frac{1}{2}$$

$$\bullet P(\{\square\bullet, \bullet\square, \square\square\square\}) = \frac{1}{2}$$

$$\bullet P(\{\square\bullet, \bullet\square\}) = \frac{1}{3}$$

What conditions must a probability function satisfy?



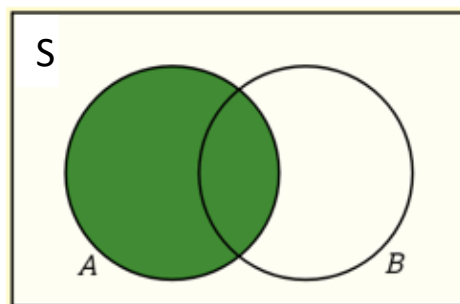
## Axioms of Probability

Axiom 1: (non-negativity)	$P(A) \geq 0$
Axiom 2: (normalization)	$P(S) = 1$
Axiom 3: (finite additivity)	If $A \cap B = \emptyset$ , Then $P(A \cup B) = P(A) + P(B)$

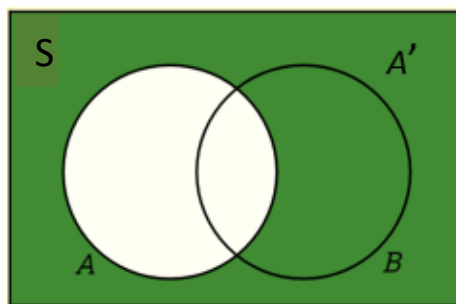
empty set

# Probability of an Event

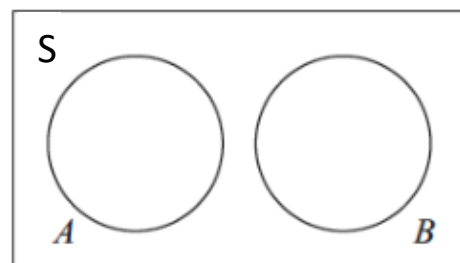
- We can define operations on events based on set theory.
- If A and B are two events:



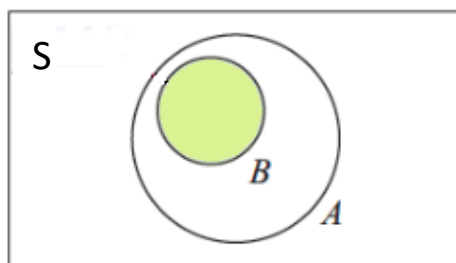
Set A



$A'$  the complement of A



A and B are disjoint sets  
(mutually exclusive)



B is proper  
subset of A  $B \subset A$

Set Operation	Venn Diagram	Interpretation
Union		$A \cup B$ , is the set of all values that are a member of A, or B, or both.
Intersection		$A \cap B$ , is the set of all values that are members of both A and B.
Difference		$A \setminus B$ , is the set of all values of A that are not members of B  (A minus B)
Symmetric Difference (parts of a union)		$A \triangle B$ , is the set of all values which are in one of the sets, but not both.

1. <https://www.datacamp.com/community/tutorials/sets-in-python>
2. <https://medium.com/@sukhrobgolibboev/understanding-set-theory-de2532f746ac>



# Useful Probability Axioms

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A') = 1 - P(A)$
- $P(S) = 1$  and  $P(\emptyset) = 0$  (something has to happen)
- $P(A \setminus B) = P(A) - P(A \cap B)$
- $0 \leq P(A) \leq 1$  for any event A
- It follows that when  $A_1, \dots, A_n$  are mutually exclusive events, then:  
$$P(A_1 \cup \dots \cup A_n) = P(A_1) + \dots + P(A_n)$$

$$P(A \cap B) = 0$$

A and B cannot occur at same time

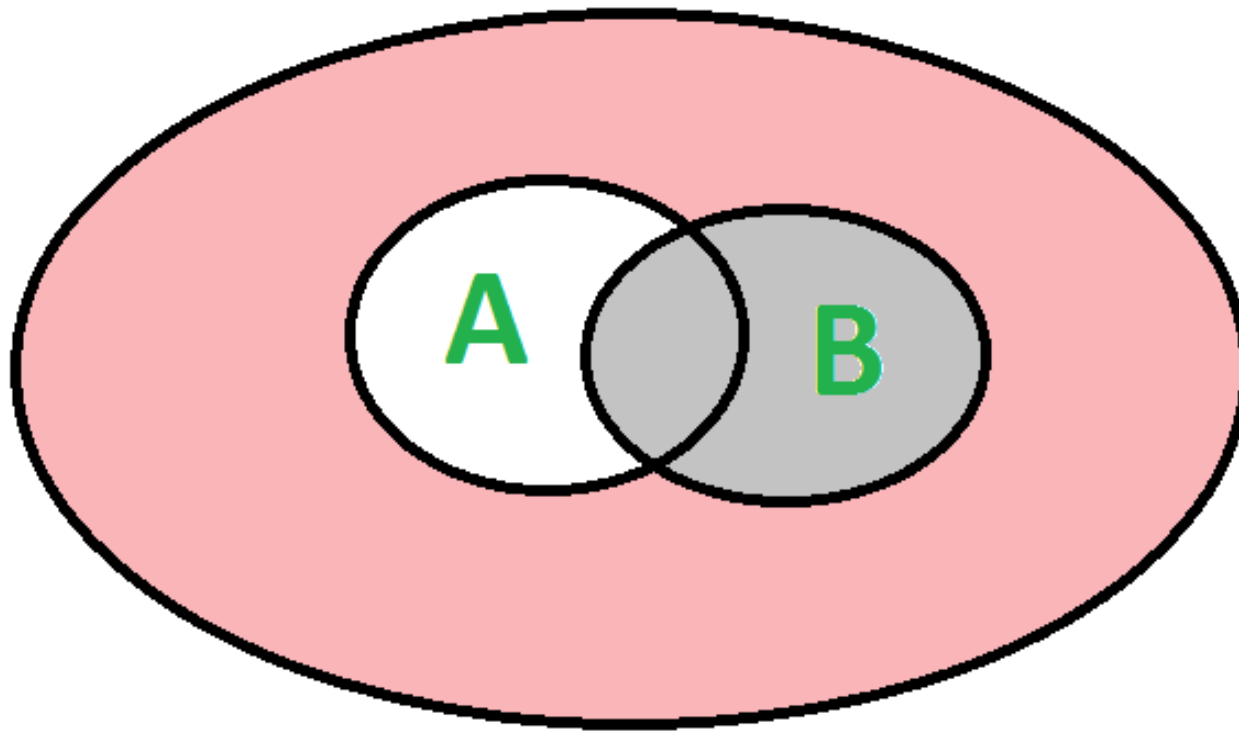


# Conditional Probability

- Conditional probability gives us a way to reason about the outcome of an experiment given such partial information.

# Conditional Probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$



Not Possible to Happen  
as B already happened

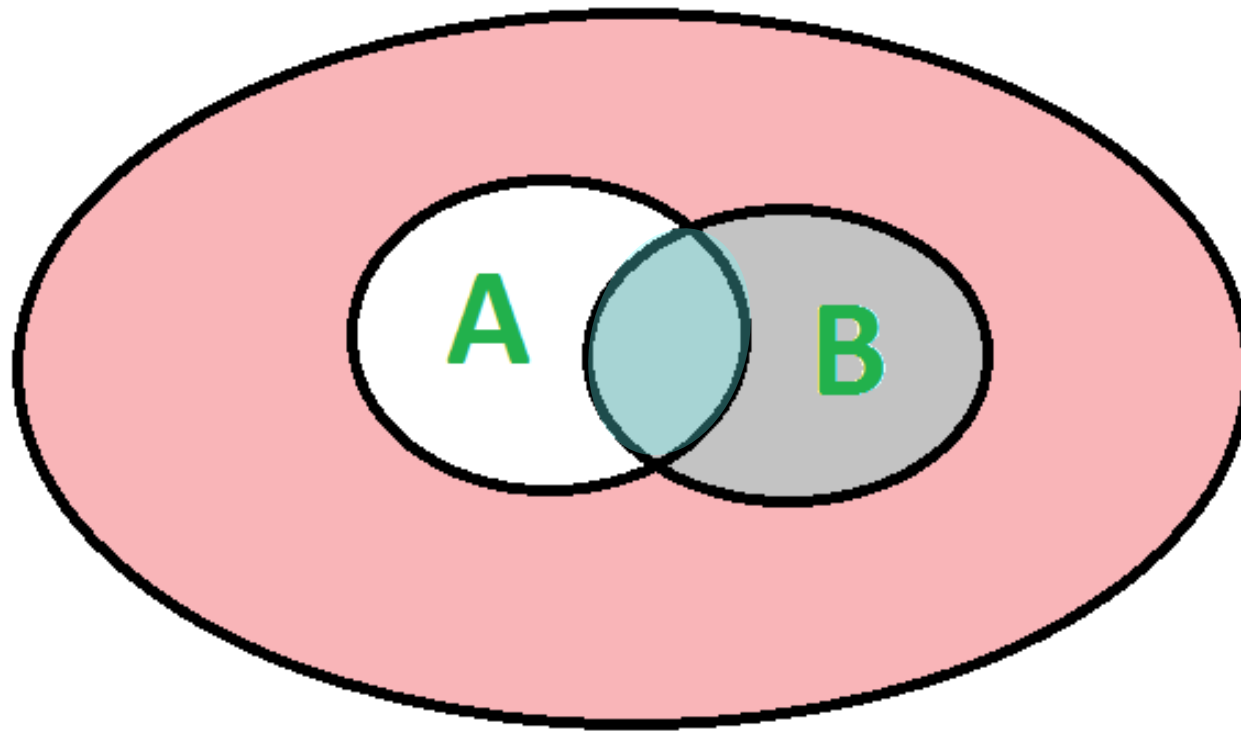


New Sample Space (After  
B happened)

Since B has already happened, the sample space reduces to B.

# Conditional Probability

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$



Not Possible to Happen  
as B already happened



New Sample Space (After  
B happened)

Since B has already happened, the sample space reduces to B.

=> Can think of this as redefining the sample space to be B, and then calculating the relative size of A within this new sample space.

# Example

- You purchase a certain product. The manual states that the lifetime  $T$  of the product, defined as the amount of time (in years) the product works properly until it breaks down, satisfies

$$P(T \geq t) = e^{\frac{-t}{5}}, \text{ for all } t \geq 0.$$

- For example, the probability that the product lasts more than (or equal to) 2 years is  $P(T \geq 2) = e^{\frac{-2}{5}} = 0.6703$ . If the product is purchased and used for two years without any problems, what is the probability that it breaks down in the third year?

# Independence

- Initial assignment of A doesn't change given that event B occurred
- A and B are independent if

$$P(A|B) = P(A) \text{ and/or } P(B|A) = P(B)$$

Otherwise, they are dependent

- When A and B are independent,

$$P(A \cap B) = P(A) P(B)$$

# General Multiplicative Law

- From the definition of the conditional probability:

$$P(A \cap B) = P(A|B)P(B)$$

$$P(A \cap B) = P(B|A)P(A)$$

- So if we extend this rule at three events, we obtain:

$$P(A \cap B \cap C) = P(A \cap B)P(C|A \cap B) = P(B|A)P(A)P(C|A \cap B)$$

- and if we extend it to n events, we obtain:

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap \cdots \cap A_{n-1})$$

# Bayes' Theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

- Bayes's theorem allows us to relate  $P(A|B)$  to  $P(B|A)$
- Provides a way to revise existing predictions or theories (update probabilities),  $P(A|B)$ , given new or additional evidence,  $P(B|A)$



The Reverend  
Thomas Bayes

Note:  $P(A|B) \neq P(B|A)$  unless  $P(A \cap B) = 0$  or  $P(A) = P(B)$



# Bayes' Theorem Physical Understanding

**Likelihood:**

How probable is the evidence  
given hypothesis is true?

**Prior:**

How probable was the hypothesis  
before observing the evidence?

$$P(H|e) = \frac{P(e|H)P(H)}{P(e)}$$

**Posterior:**

How probable is the hypothesis  
given the observed evidence?

**Marginal:**

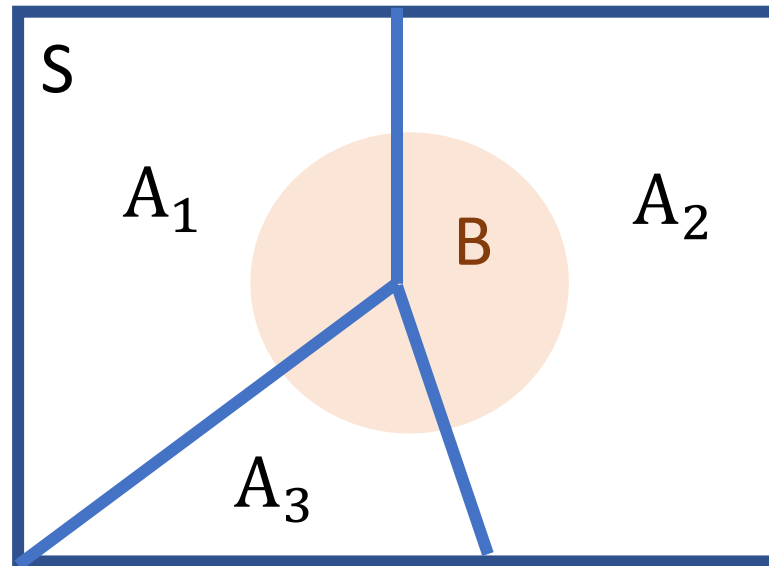
How probable is the new evidence  
under all possible hypothesis?

# Law of Total Probabilities

- $P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + \cdots + P(B|A_n)P(A_n)$

where  $A_1, \dots, A_n$  is a partition of the sample space

$S(A_i \cap A_j = \emptyset \text{ for any } i \neq j \text{ and } A_1 \cup A_2 \cup \cdots \cup A_n = S)$



# Bayes' Theorem

- Applying Law of Total Probabilities to Bayes' Theorem

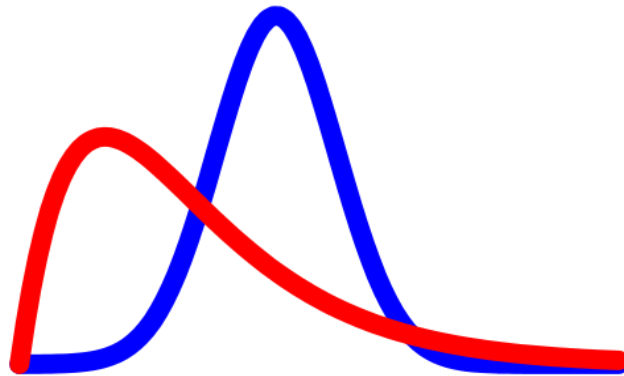
$$\begin{aligned} P(H|e) &= \frac{P(e|H)P(H)}{P(e)} \\ &= \frac{P(e|H)P(H)}{P(H)P(e|H) + P(\neg H)P(e|\neg H)} \end{aligned}$$

  $\neg$  means "not"

# Bayes' Theorem

- Given a priori probabilities (assignments)  $P(A_1), P(A_2), \dots, P(A_n)$ , (where  $A_1; \dots, A_n$  is a partition of the sample space  $S$ ), we can compute (**update**) the posterior probabilities  $P(A_i|B)$ , for  $i = 1, \dots, n$  based on the conditional probabilities  $P(B|A_i)$ , for  $i = 1, \dots, n$  using:

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{k=1}^n P(B|A_k) P(A_k)}$$

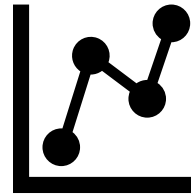


# When to use Bayes' Theorem

You have a hypothesis



You've observed evidence



You want

**$P(\text{hypothesis} \text{ given evidence})$**

- Quantify and systematize the idea of changing beliefs or updating info
  - Science: analyzing extent to which new data validates or invalidates model
  - AI: explicitly and numerically model a machine's belief
  - Reframes how you think about thought itself

See: 3blue1brown for more intuitive info

<https://www.youtube.com/watch?v=HZGCoVF3YvM&t=206s>

# Spam Mail Example

- It is estimated that 50% of emails are spam emails. Some software has been applied to filter these spam emails before they reach your inbox. A certain brand of software claims that it can detect 99% of spam emails, and the probability for a false positive (a non-spam email detected as spam) is 5%. Now if an email is detected as spam, then what is the probability that it is in fact a non-spam email?

# Random Variables

- A random variable,  $RV$ , is a mapping from the sample space  $S$  to the real line:

$$RV: S \rightarrow \mathbb{R}$$

- i.e., assigns a real number to every possible outcome in the sample space, where  $S$  is the sample space (the space of all possible outcomes) and  $\mathbb{R}$  is called state space or range space.
- Application:
  - In a complex system
    - the number of days until a part failure
    - the number of parts that have failed today
    - the number of customers affected by a failure

# Notation

- The state space consists of the associated values of the outcomes in the sample space according to the rule  $RV$ :

$$RV(o_i) = x_i$$

through which we assign a numerical value  $x_i$  to each outcome  $o_i$  .

- Can also be written as:

$$P(\{X = k\}) \text{ for } k = 2,3,4$$

- Capital letters: denote random variables
- Lowercase letters: denote particular outcomes a random variable might take

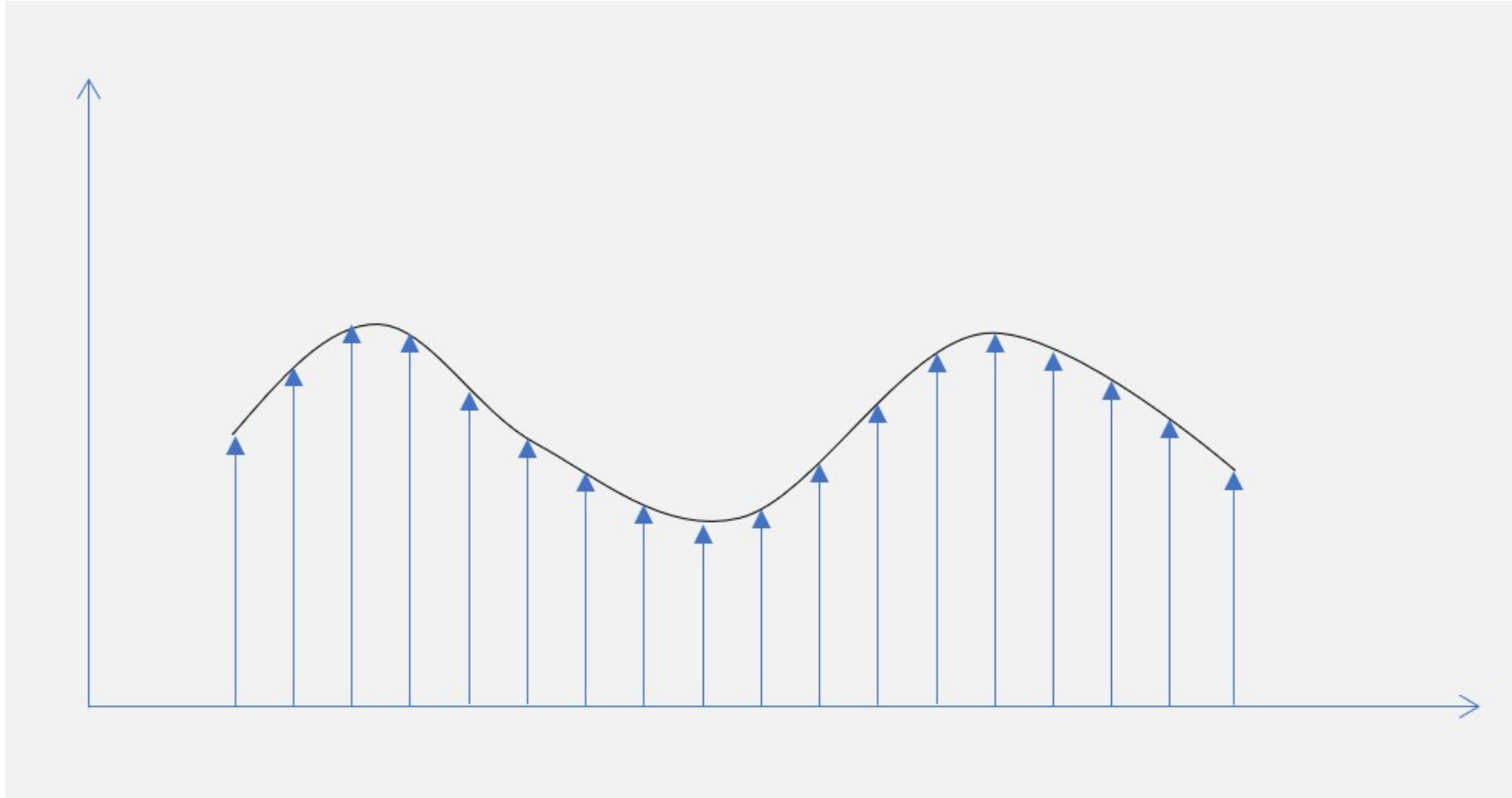


# Example

- In any experiment there are various characteristics that can be observed or measured. In most of the studies, an experimenter will focus on some aspect of the experiment.
- A researcher may test a sample of components from a production line and record only the number of components that have failed within 100 hours. In this example, for each component we observe, 0 (failed) or 1 (not failed).
- If observe one outcome to be 100 of 1's (none of the components failed), call it  $o_{100}$ . Another outcome may be observing 90 of 1's (only 10 of the components failed), call it  $o_{90}$ , and so on. Here one random variable associates the frequency value 100 to  $o_{100}$  and the frequency value 90 to  $o_{90}$ :

$$RV(o_{100}) = 100, \quad RV(o_{90}) = 90$$

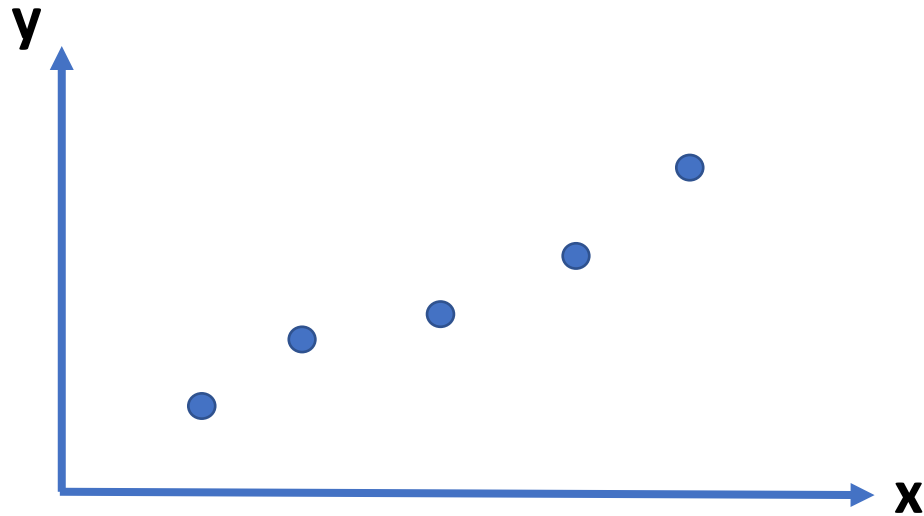
# Discrete vs. Continuous Numbers



<https://www.sigmamagic.com/blogs/is-my-data-continuous-or-discrete/>

# Discrete Random Variables

- For a discrete *RV*, the state space  $\mathcal{R}$  is discrete or countable.
- Example of discrete space is  $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$  and an example of a countable space is the set of all integers  $\mathbb{Z}$ .



# Probability Mass Function (PMF)

- For a discrete rv  $X$ , we define the probability mass function to be a function:

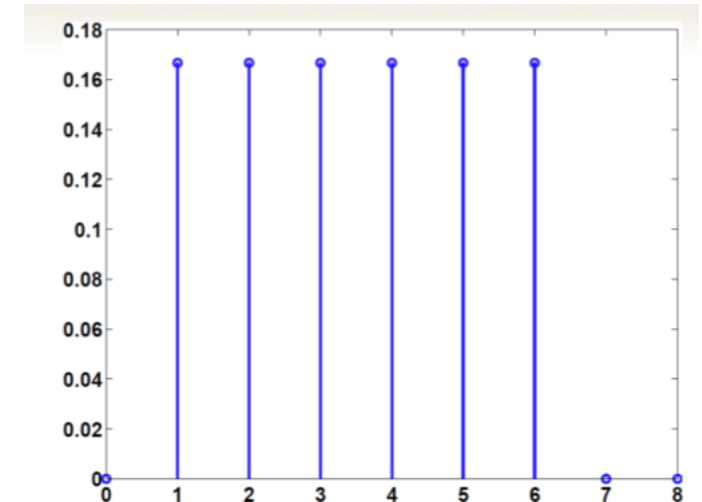
$$P_x: \mathcal{R} \rightarrow [0,1]$$

for a state space  $\mathcal{R} = \{x_1, x_2, \dots, x_n\}$  such that

$$P_x(x_i) = p_i, \quad i = 1, \dots, n$$

- The values  $p_i$  for  $i = 1, \dots, n$  are such that  $0 \leq p_i \leq 1$  and  $\sum_{i=1}^n p_i = 1$
- This is a simple function that tells us the probability of each possible outcome
- Used for **discrete** probability distributions

Roll of a fair die



$$p_X(k) = \begin{cases} \frac{1}{6} & k = 1, 2, \dots, 6 \\ 0 & \text{otherwise} \end{cases}$$



# Cumulative Distribution (CDF for discrete RV)

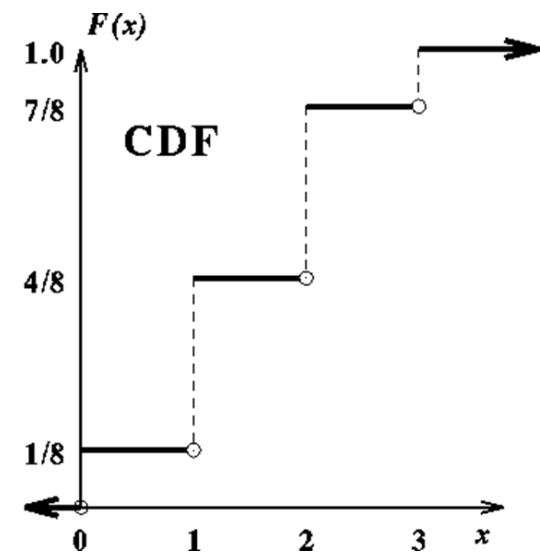
- For a discrete rv  $X$ , we define the cumulative distribution function (cdf) a function:

$$F_x: \mathcal{R} \rightarrow [0,1]$$

for a state space  $\mathcal{R} = \{x_1, x_2, \dots, x_n\}$  such that

$$F_x(x) = P_x(X \leq x) = \sum_{x_i \leq x} P_x(X = x_i)$$

- In other words, the cdf is the probability that a random variable is less than or equal to a certain real number



# Mean

- For a discrete rv  $X$  with a state space  $\mathcal{R} = \{x_1, x_2, \dots, x_n\}$  and pmf  $P_X$ , we define its expectation (expected value) or mean to be a weighted average of the values in the state space:

$$E(X) = \sum_{i=1}^n x_i P(X = x_i)$$

- The expectation is a measure for the average value in the state space
  - a “weighted average” over all the values  $X$  can take, where the weights are given by the probabilities of each of those values
- Note: expectation is not random
  - it is a deterministic function of  $P(X = x_i)$
- Note: expectation does not need to be one of the possible outcomes

# Example

- Apple is interested in hiring you to do some consulting for them. They have two projects they would like your help with, but they will only pay you if they are satisfied with your work.
  - Project 1 pays \$1000 and you believe that the probability you will complete the project to Apple's satisfaction is 0.8.
  - Project 2 pays \$2000 and you believe that the probability you will complete the project to Apple's satisfaction is 0.5.
- If Apple is happy with your work on whichever project you choose to do first, they will give you the chance to do the second project, but if they don't like your work they will send you on your way. Which project should you take first to maximize your expected earnings?

# Variance

- Expectation tells us about the average outcome, but also often need to know how likely rv is close to the average outcome
- For a discrete  $X$  with a state space  $\mathcal{R} = \{x_1, x_2, \dots, x_n\}$  and pmf  $P_X$ , we define its variance to be:

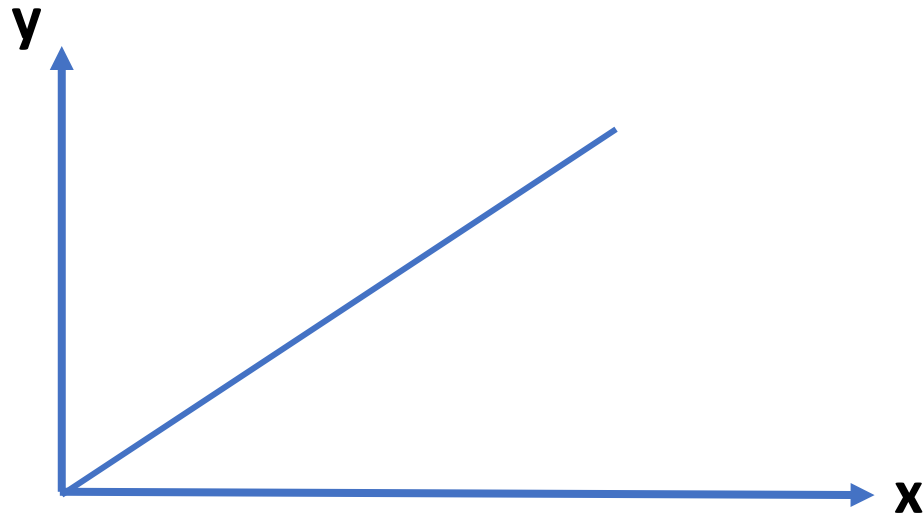
$$\begin{aligned} V(X) &= \sum_{i=1}^n (x_i - E(X))^2 P(X = x_i) \\ &= \sum_{i=1}^n x_i^2 P(X = x_i) - (\sum_{i=1}^n x_i P(X = x_i))^2 \end{aligned}$$

- The variance is a measure for the variability or spread in the state space.
- Describes how much a random variable differs from its expected value.
- The standard deviation of  $X$  is  $\sqrt{V(X)}$ 
  - Standard deviation is often easier to interpret since it has the same units as  $X$



# Continuous Random Variables

- For a continuous rv, the state space  $\mathcal{R}$  is infinite. Examples of infinite spaces are  $[0, 1]$  and the set of all real numbers.



# Probability Density Function (PDF)

- For a continuous rv  $X$  for a state space  $\mathcal{R}$ , we define its probability density function (pdf) to be a function:

$$f_x: \mathcal{R} \rightarrow \mathbb{R}$$

such that

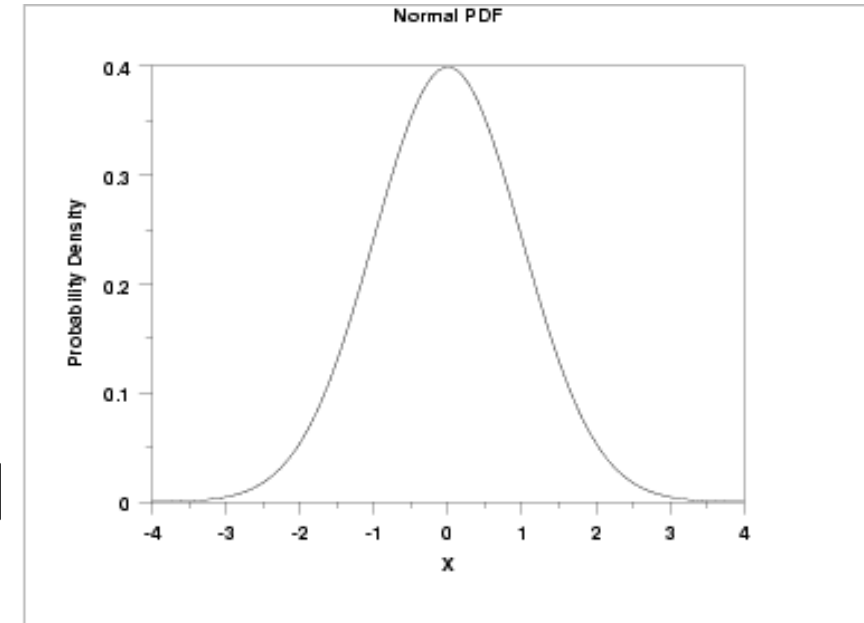
$$(1) f(x) \geq 0$$

$$(2) \int f(x)dx = 1$$

- We define the density area above an interval  $[a, b]$  for  $a \leq b$  to be:

$$P(a \leq X \leq b) = \int_a^b f_x(x)dx$$

- Used for continuous probability distributions



<https://www.itl.nist.gov/div898/handbook/eda/section3/eda362.htm>

# Cumulative Distribution Function (CDF for continuous RV)

- For a continuous rv  $X$  with a state space  $\mathcal{R}$ , we define its cumulative distribution function (cdf) a function:

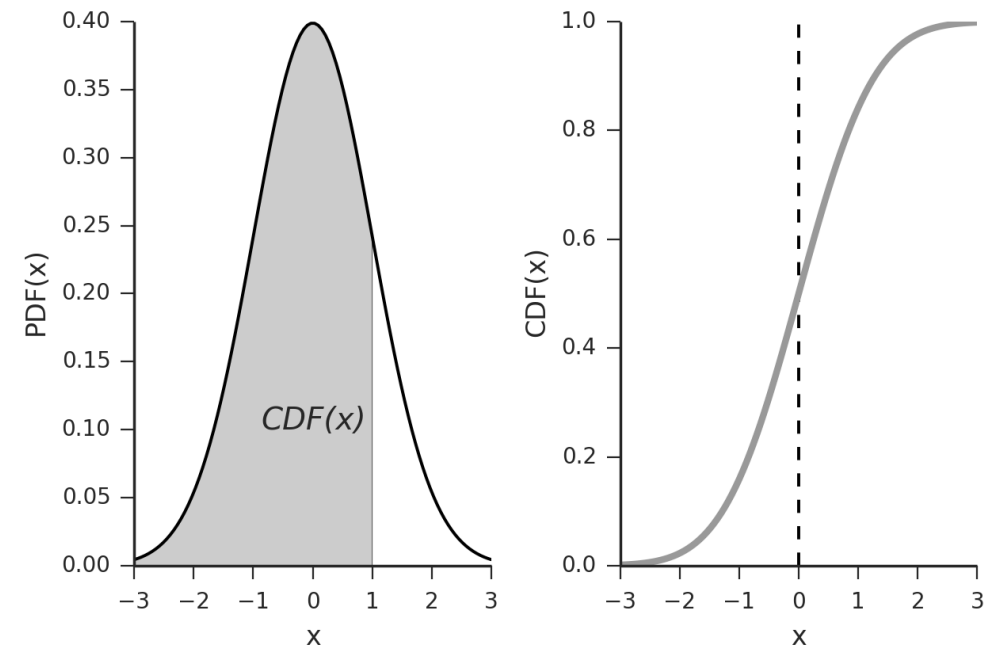
$$F_x: \mathcal{R} \rightarrow [0,1]$$

such that

$$F_x(x) = P(X \leq x) = \int_{-\infty}^x f_x(x) dx$$

- One important property of the cdf is that its first derivative is:

$$F'_x(x) = f_x(x)$$



# Mean

- For a continuous rv  $X$  with a state space  $R$  and pdf  $f_x$ , we define its expectation or mean to be:

$$E(X) = \int x f_x(x) dx$$

The expectation is a measure for the average value in the state space.

# Example

- The time to failure in thousands of hours of an important piece of electronic equipment used in a manufactured radio has the density function

$$f(x) = \begin{cases} 3e^{-3x}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

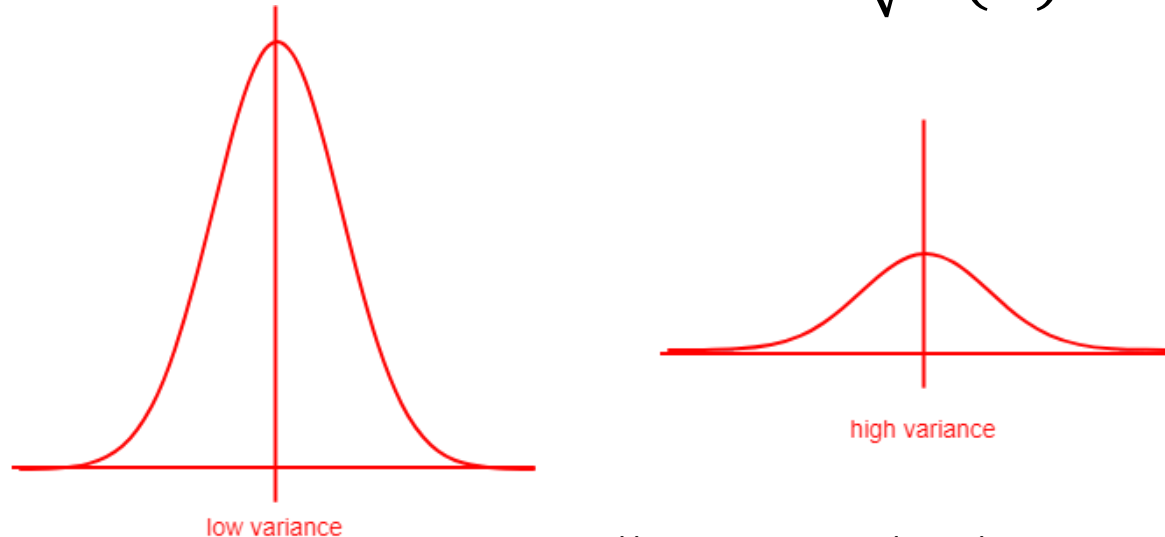
- Find the expected life of the piece of equipment.

# Variance

- For a continuous rv  $X$  with a state space  $R$  and pdf  $f(x)$ , we define its variance to be:

$$V(X) = \int (x - E(X))^2 f(x) dx = \int x^2 f(x) dx - \left( \int x f(x) dx \right)^2$$

- The variance is a measure for the variability or spread about the mean in the state space. The standard deviation of  $X$  is  $\sqrt{V(X)}$ .



# Properties of Mean and Variance

- Useful properties for a random variable  $X$ , with constants  $a, b \in \mathbb{R}$ :
  - $E[X + b] = E[X] + b$
  - $E[aX] = aE[X]$
- Can combine above two properties into one statement:
  - $E[aX + b] = aE[X] + b$

# Properties of Mean and Variance

- One important property for variance for both discrete and continuous distributions is:

$$V(X) = E(X^2) - (E(X))^2$$

- Proof:
- $V(X) = E[(X - E[X])^2]$  (original definition of variance)  
$$= E[(X^2 - 2XE(X)) + (E[X])^2]$$
$$= E(X^2) - 2E(X)E(X) + (E(X))^2$$
$$= E(X^2) - (E(X))^2$$



# Properties of Mean and Variance

- $\text{Var}(X + c) = \text{Var}(X)$

- Proof:

$$\begin{aligned}\text{Var}(X + c) &= E[(X + c)^2] - E(X + c)^2 \\ &= E(X^2 + 2cX + c^2) - E(X + c)E(X + c)\end{aligned}$$

Expanding the first term,

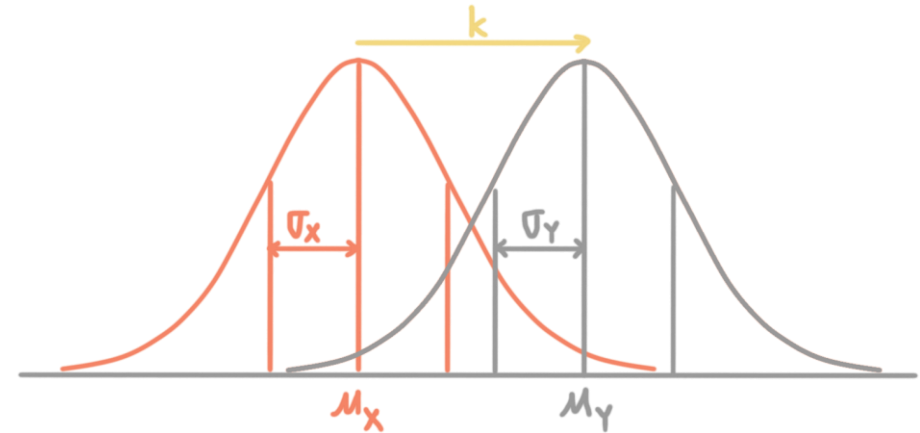
$$E(X^2 + 2cX + c^2) = E(X^2) + 2cE(X) + c^2$$

Expanding the second term,

$$\begin{aligned}E(X + c)E(X + c) &= E(X)E(X + c) + E(c)E(X + c) \\ &= E[XE(X) + cE(X)] + cE(X) + cE(c) \\ &= E(X)^2 + cE(X) + cE(X) + c^2 \\ &= E(X)^2 + 2cE(X) + c^2\end{aligned}$$

Putting it all together,

$$\begin{aligned}\text{Var}(X + c) &= E(X^2) + 2cE(X) + c^2 - E(X)^2 - 2cE(X) - c^2 \\ &= E(X^2) - E(X)^2 \\ &= \text{Var}(X)\end{aligned}$$



- Since the variance measures the amount of spread of the distribution, shifting the distribution left or right by a constant doesn't affect that spread and therefore shouldn't affect the variance.

# Properties of Mean and Variance

- $Var(cX) = c^2 var(X)$
- Proof:

$$\begin{aligned} Var(cX) &= E[(cX)^2] - E(cX)^2 \\ &= c^2 E(X^2) - c^2 E(X)^2 \\ &= c^2 [E(X^2) - E(X)^2] \\ &= c^2 Var(X) \end{aligned}$$

# Example

- A fisherman is weighing each of 50 fishes. Their mean weight worked out is 50 gm and a standard deviation of 2.5 gm. Later it was found that the measuring scale was misaligned and always under reported every fish weight by 2.5 gm. Find the mean and standard deviation of fishes.

# Moments of a Distribution

- We define the  $k^{th}$  moment of a distribution  $X$  (for both discrete and continuous distributions) to be the expectation of  $X^k$ .

- For a discrete distribution, the  $k^{th}$  moment is

$$E(X^k) = \sum_{x \in \mathcal{R}} x^k P(X = x)$$

- For a continuous distribution, the  $k^{th}$  moment is

$$E(X^k) = \int x^k f_x(x) dx$$

# Combination and functions of random variables

- If  $Y$  is a random variable such that it can be expressed in the form:

$$Y = aX + b$$

where  $X$  is a random variable then we have:

$$E(Y) = aE(X) + b$$

$$V(Y) = a^2V(X)$$

- For a sequence of rv's  $X_1, X_2, \dots, X_n$ :

$$E(a_1X_1 + \dots + a_nX_n) = a_1E(X_1) + \dots + a_nE(X_n)$$

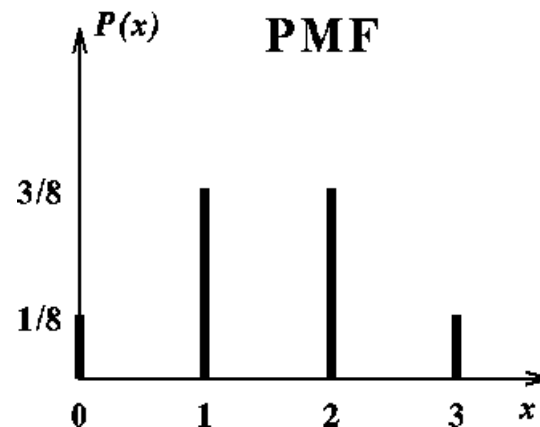
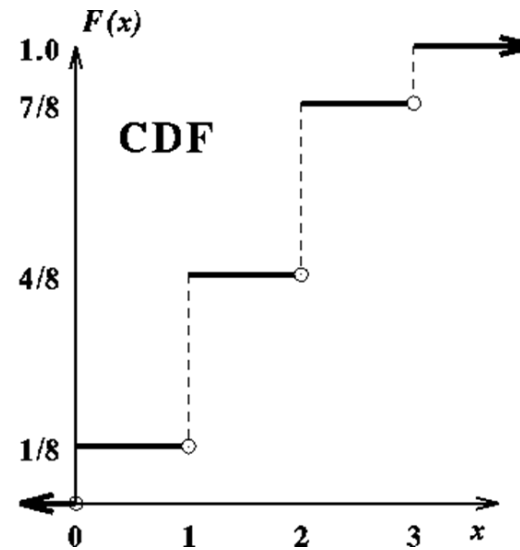
# Summary

- Independence:  $P(A \cap B) = P(A) P(B)$
- Conditional Probability:  $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- Bayes' Theorem :  $P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(A \cap B)}{P(B)}$ 
  - Used to update pre-existing condition
  - $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$

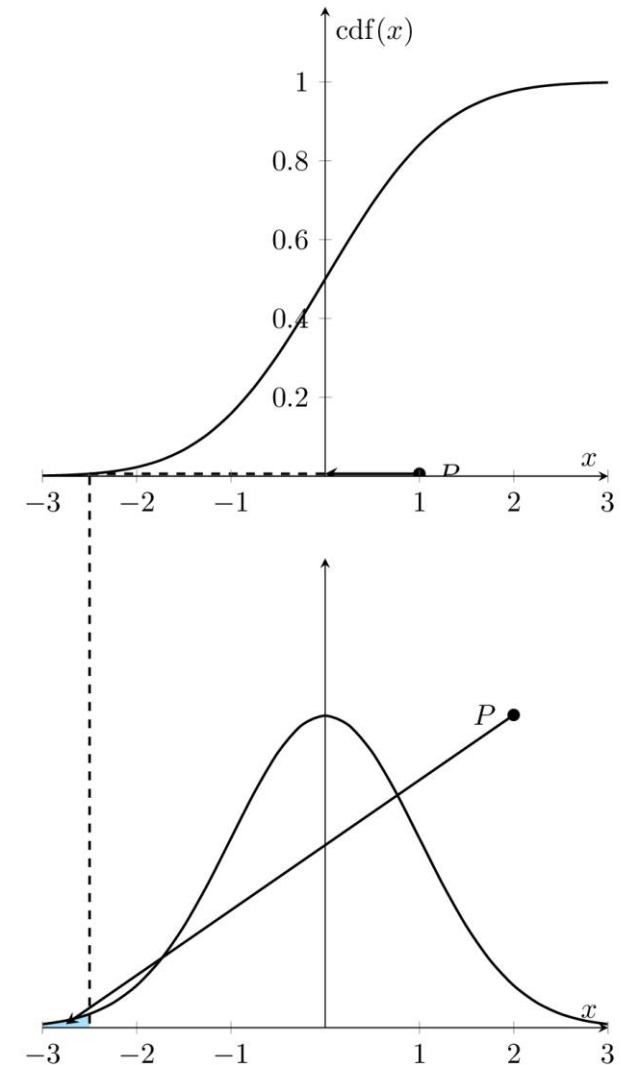
# Summary

- Cumulative distribution function (CDF): probability that a random variable is less than or equal to a certain real number
- Probability mass function (PMF): probability that a discrete random variable equals a specific value or density for a discrete random variable
- Probability distribution function (PDF): describes the probability distribution of a continuous random variable (relative probability)

## Discrete Random Variable



## Continuous Random Variable



# Summary

- Useful equations:
  - $Var(X) = E(X^2) - (E(X))^2$
  - $E[aX + b] = aE[X] + b$
  - $Var(X + c) = Var(X)$
  - $Var(cX) = c^2 var(X)$