1 Probability Space

There are often various approaches to probability each with its own advantages and disadvantages.

Experiment is a procedure that can be infinitely repeated and has a well-defined set of possible outcomes, known as the sample space.

The observation/result of the experiment are termed as outcomes.

1.1 Classical Approach

Probability of an event E is defined to be:

$$P(E) = \frac{\text{Number of outcomes in } E}{\text{Total number of outcomes}}$$

Some examples are tossing a coin or rolling a die. Disadvantages:

- Unable to model biases. It says nothing about cases where no physical symmetry exists.
- Doesn't deal with cases where total outcomes are infinite.

1.2 Frequentist Approach

Also known as the relative frequency approach or frequentism. It defines an event's probability as the limit of its relative frequency in many trials.

Probability is defined to be:

$$P(E) = \lim_{n \to \infty} \frac{n_E}{n}$$

where an experiment is conducted n times and event E occurs n_E times. Disadvantages:

- It isn't efficient to conduct an experiment multiple times just to find the probability of an event occuring.
- It is unable to deal with subjective belief. Eg: Suppose a cricket expert says there is a 50% of RCB winning the IPL this year. It doesn't mean that the RCB has won half the titles in the past.

1.3 Axiomatic Approach

1.3.1 Probability Space

The triple (S, F, P) is referred to as a probability space where:

- S : Sample space, set of all possible outcomes of the experiment.
- F : Event Space, subset of the sample space
- P : Probability Measure

1.3.2 Sample Space

S can either be finite or countably infinite or uncountably infinite.

Examples for S:

- Finite Sample Space: Single coin toss $S = \{H, T\}$ and two coin tosses $S = \{HH, HT, TT, TH\}$.
- Countably Infinite Sample Space: Keep tossing a coin till you get a head $S = \{H, TH, TTH...\}$.
- Uncountably Infinite Sample Space: We have a circular dart board and we are measuring the angle at which a dart hits the board. $S = [0, 2\pi]$

1.4 Event Space

Collection of events is called an event space, there are some properties to be satisified such as: it has to be a "Sigma Field".

1.4.1 Sigma Field

A sigma field (or sigma algebra) F is a collection of subsets of S which satisfies the following properties:

- $S \in F$
- If $E \in F$, then $E^C \in F$
- If $E_1, E_2, E_3 \cdots \in F$, then $\bigcup_{i=1}^{\infty} E_i \in F$

1.4.2 Examples for Event Space

- Smallest possible event space:

$$F = \{\phi, S\}$$

- Next non-trivial event space:

$$F = \{\phi, E, E^c, S\}$$

- If $E_1 \in F$ and $E_2 \in F$, then $E_1 \cap E_2 \in F$ (Proof in 1.4.3)
- For S=$\{1, 2, 3, 4, 5, 6\}$, $E_1 = 1, 2$ and $E_2 = 3, 4$. The smallest event space containing E_1 and E_2 is:

$$F = \{\phi, S, E_1, E_1^C, E_2, E_2^C, E_1 \cup E_2, (E_1 \cup E_2)^C\}$$

1.4.3 Proposition 1

$A_1, A_2, A_3, \ldots A_n \in F$, then $\bigcap_{i=1}^{n} A_i \in F$.

Proof: If $A_1, A_2, A_3, \ldots A_n \in F$, then $A_1^c, A_2^c, A_3^c, \ldots A_n^c \in F$ and $\bigcup_{i=1}^{n} A_i^c \in F$. Then by property 2, $(\bigcup_{i=1}^{n} A_i^c)^c = \bigcap_{i=1}^{n} A_i \in F$.

1.4.4 Proposition 2

A, B $\in F$, then A\B = A - B $\in F$.

Proof: If B $\in F$, then $B^c \in F$ by property 2. So, $A \cap B^c$ = A\B $\in F$ (As seen in proposition 1).

1.5 Probability Measure

The probability measure P is a function returning an event's probability. A probability is a real number between zero and one.

$$P : F \to [0, 1]$$

P has to satisfy the following 3 axioms:

- $P(E) \geq 0$
- $P(S) = 1$
- If $E_1, E_2 \cdots \in F$ such that $E_i \cap E_j = \phi$ then:

$$P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$$

For two disjoint sets: $P(E_1 \cup E_2) = P(E_1) + P(E_2) + \sum P(\phi)$
We will later see that $P(\phi)$ is indeed 0.

1.6 Derived Properties of Probability

1. $$P(E^C) = 1 - P(E)$$

   Proof:

$$E \cup E^C = S$$

$$P(E) + P(E^C) = 1$$

2. For any two events E_1 and E_2,

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$

Proof:

$$P(E_1 \cup E_2) = P(E_1) + P(E_2 \cap E_1^C)$$

Now,

$$E_2 = (E_2 \cap E_1) \cup (E_2 \cap E_1^C)$$

$$\Rightarrow P(E_2) = P(E_1 \cap E_2) + P(E_1^C \cap E_2)$$

Also,

$$P(E_1 \cup E_2) = P(E_1) + P(E_1^C \cap E_2)$$

Substitute the required value in the final equation.

Question: $S = \{1, 2, 3, 4, 5, 6\}$. 1 and 5 are equally likely and probability of getting a 6 is one-third.

Find minimum and maximum probability that we get an even number.

Answer:

$$Minimum\ Prob. = \frac{1}{3}\ \ when\ \ P_2 = P_4 = 0$$

$$Maximum\ Prob. = 1,\ \ P_2 + P_4 = \frac{2}{3}$$

2 Conditional Property

Given that an event A has occured.

$(S, F, P) \rightarrow$ Original probability space

If additional info has been given that A has occured, probability space need to be suitably modified.

eg: $S = \{1, 2, 3, 4, 5, 6\},\ E_1 = \{1, 2\},\ E_2 = \{3, 4\}$

$$F = \{\phi, S, E_1, E_1^C, E_2, E_2^C, E_1 \cup E_2, (E_1 \cup E_2)^C\}$$

and event $A = E_1^c = \{3, 4, 5, 6\}$ has occured.

Then, $F_A = \{\phi, A, \{3, 4\}, \{5, 6\}\}$.(shown later)

2.1 Modified probability space

Then,

- $S_A = A$. (Modified Sample Space)
- $F_A = \{(E \cap A) | E \in F\} \rightarrow E \cap A \in F$. (Modified Event Space)

  (Also if some event $C \cap A = \phi$, then C won't occur)

  To prove: F_A also satisfies event space axioms.(see sec. 1.4.1)

1. $A \in F_A \rightarrow S \cap A = A$(S was original sample space) $\Rightarrow A \in F_A$
2. $D \in F_A \Rightarrow D = E \cap A, D^c \in F_A$
   As,
   $D^c = A \setminus D = E^c \cap A \in F \quad (Because\ E^c \in F)$
3. $$D_1, D_2, \cdots \in F_A$$
   $$(E_1 \cap A), \cdots \in F_A$$
   $$E_1, E_2 \cdots \in F$$
   $$\Rightarrow \bigcup_{i=1}^{\infty} E_i \in F_A$$
   $$\Rightarrow (\bigcup_{i=1}^{\infty} E_i) \cap A \in F_A$$
   $$\Rightarrow (\bigcup_{i=1}^{\infty} E_i \cap A) \in F_A \Rightarrow \bigcup_{i=1}^{\infty} D_i \in F_A$$

Hence, F_A is an event space.

- Modified probability measure

$$P(E/A) = \frac{P(E \cap A)}{P(A)}$$

This definition is called conditional probability measure for any $E \in F$.

eg: $F_A = \{\phi, \{3,4,5,6\}, \{3,4\}, \{5,6\}\}$, then $P(\{3,4\}/\{3,4,5,6\}) = 1/2$

Now, we need to prove $P(E/A)$ satisfies the 3 axioms of probability measure.(see sec. 1.5)

  - $P(E/A) \geq 0$ (as ratio of two nos. which are positive)
  - $P(S/A) = 1$
  - $B_1, B_2 \cdots$ are all mutually disjoint.

$$P(\bigcap_{i=1}^{\infty} B_i/A) = \frac{P(\bigcup_{i=1}^{\infty} B_i \cap A)}{P(A)}$$

$$= \frac{\sum_{i=1}^{\infty} P(B_i \cap A)}{P(A)} = \sum_{i=1}^{\infty} P(B_i/A)$$

$$\Rightarrow P(\bigcap_{i=1}^{\infty} B_i/A) = \sum_{i=1}^{\infty} P(B_i/A)$$

3 Total probability theorem

Events $A_1, \cdots, A_n \in F$ which are all mutually exclusive/disjoint and exhaustive. Then,

$$A_i \cap A_j = \phi \quad \forall\, i, j$$

$$\bigcup_{i=1}^{n} A_i = S$$

$$P(B) = \sum_{i=1}^{n} P(B/A_i)P(A_i)$$

Expresses $P(B)$ in terms of conditional probability $P(B/A_i)$ & prior probability $P(A_i)$

Proof:

$$B = \bigcup_{i=1}^{n} (B \cap A_i)$$

A_i's are disjoint, so $(B \cap A_i)$ are also disjoint.

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B/A_i)P(A_i)$$

3.1 Question

Two factories manufacture zoggles. 20% of F_1 are defective. 5% of F_2 are defective.

In any week, F_1 produces twice the number of zoggles as F_2. What is the probability that a zoggle chosen randomly in a week is defective?

$$\begin{aligned} P(D) &= P(F_1)P(D/F_1) + P(F_2)P(D/F_2) \\ &= \frac{2}{3} \times \frac{1}{5} + \frac{1}{3} \times \frac{1}{20} = \frac{3}{20} \end{aligned}$$

4 Bayes Theorem

$A_1, \cdots, A_n$ are events which are mutually exclusive and exhaustive. B be an arbitrary event. Then,

$$P(A_i/B) = \frac{P(B/A_i)P(A_i)}{\sum_{i=1}^{n} P(B/A_i)P(A_i)}$$

- $P(A_i)$- Prior probability
- $P(B/A_i)$- Likelihood
- $P(A_i/B)$- Posterior probability

Bayes theorem expresses posterior probabilities $P(A_i/B)$ in terms of prior probabilities $P(A_i)$ and likehoods $P(B/A_i)$.

In some experiments $P(A_i)$ are all same, $P(A_i) = 1/n$. Then posterior probabilities are proportional to likelyhoods:

$$P(A_i/B) = \frac{P(B/A_i)}{\sum_{i=1}^{n} P(B/A_i)}$$

$$\Rightarrow P(A_i/B) \propto P(B/A_i)$$

Proof:

$$P(A_i/B) = \frac{P(A_i \cap B)}{P(B)}$$

$$= \frac{P(B/A_i)P(A_i)}{\sum_{i=1}^{n} P(B/A_i)P(A_i)}$$

4.0.1 Example

In answering a question in a multiple choice test, a student knows the answer with probability p and guesses the answer otherwise. If he/she guesses from m choices, the probability of being correct is $\frac{1}{m}$. Find the conditional probability that the student knew the answer if he/ she answered correctly.

Ans: A_1 = knowing the answer. A_2 = Guessing the answer.

B =Answer is correct. $A_1 \subseteq B$

$P(A_1) = p, P(A_2) = 1 - p, P(B/A_1) = 1, P(B/A_2) = \frac{1}{m}$

$$P(A_1/B) = \frac{P(B/A_1)P(A_1)}{P(B/A_1)P(A_1) + P(B/A_2)P(A_2)}$$

$$= \frac{p}{p + (1-p)\frac{1}{m}}$$

5 Indepedent Events

A & B are said to be independent. If

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

In terms of conditional probability,

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

$$\Rightarrow P(B/A) = P(B)$$

Probability of event B remains same with or without conditioning on A. Hence, B is said to be independent of A. Knowledge of occurence of event A does not give any information about B.

If A & B are independent, then A & B^c are also independent.
Proof:

$$P(B^c/A) = \frac{P(B^c \cap A)}{P(A)} = \frac{P(A) - P(A \cap B)}{P(A)} = \frac{P(A) - P(A)P(B)}{P(A)} = P(B^c)$$

5.1 Important Results

1. $P(A/B)$ may be greater than, less than or equal to $P(A)$.

2. Independent events and mututally exclusive events are different.

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

   Mututally exclusive: $A \cap B = \phi$

   Eg 1: (Indepedent but not mutually exclusive). Coin toss followed by throwing dice experiment.

   $$S = \{(H,1),(H,2),\cdots,(H,6),(T,1),\cdots,(T,6)\}$$

   F = Power set of S (Always an event space)

   $A = \{(H,1),\cdots,(H,6)\}$, $B = \{(H,2),(H,4),(H,6),(T,2),(T,4),(T,6)$

   $P(A) = 1/2$ & $P(B) = 1/2$

   $$P(A \cap B) = P(\{(H,2),(H,4),(H,6)\}) = \frac{3}{12} = \frac{1}{4}$$

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

   Eg 2: (Not independent but mutually exclusive) If the events are mutually exclusive $\Rightarrow$ they are not independent.

   Single Coin Toss: $A = \{H\}$, $B = \{T\} \rightarrow$ Mututally exclusive.

   $P(A \cap B) = 0$, $P(A)P(B) = \frac{1}{4}$

5.2 Conditionally indepedent events

A & B are said to be conditionally indepedent given C if

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

In terms of conditional probabilities,

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

Indepedent Events: $P(A \cap B) = P(A)P(B)$ & $P(B/A) = P(B)$.
Conditionally indepedent events: $P(A \cap B/C) = P(A/C)P(B/C)$ & $P(B/(A \cap))$

5.2.1 Example

Two fair coins are tossed, $S = \{HH, HT, TH, TT\}$, $A = \{HH, HT\}$, $B = \{HH, TH\}$, $C = \{HH\}$ & $D = \{HT, TH\}$

1. Are A & B independent?

$$P(A \cap B) = P(\{HH\}) = \frac{1}{4} = P(A)P(B)$$

2. Are A & B conditionally indepedent given C?

$$P((A \cap B)/C) = P(A/C) = P(B/C) = 1$$

$$P(A/C) = \frac{P(A \cap C)}{P(C)} = \frac{\frac{1}{4}}{\frac{1}{4}} = 1$$

3. Are A & B conditionally indepedent given D?

$$P(A/D) = \frac{P(A \cap D)}{P(D)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$$

$$P(B/D) = \frac{P(B \cap D)}{P(D)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$$

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

Hence, they are not conditionally indepedent given D.

If A and B are independent, it doesn't imply A and B will be conditionally indepedent given C and vice-versa.

5.3 Indepedence of collection of events

Three events A_1, A_2 & A_3 are said to be indepedent if:

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$$

$$P(A_1 \cap A_2) = P(A_1)P(A_2) \quad P(A_2 \cap A_3) = P(A_2)P(A_3) \quad P(A_3 \cap A_1) = P(A_3)P(A_1)$$

5.3.1 Chain rule of Probability

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2/A_1)P(A_3/A_1 \cap A_2)$$

$$P(\bigcap_{i=1}^{n} A_i) = P(A_1)\prod_{i=2}^{n} P(A_i/A_1 \cdots A_{i-1})$$

For independent events A_1, A_2, A_3.

$$P(A_1)P(A_2/A_1)P(A_3/(A_1 \cap A_2)) = P(A_1)P(A_2)P(A_3)$$

$$\Rightarrow P(A_2/A_1)P(A_3/(A_1 \cap A_2)) = P(A_2)P(A_3)$$

Hence, if A_1, A_2, A_3 are independent events, the first condition doesn't imply the other conditions.

Eg: Pair-wise independence does not imply that three events are independent. Consider example of two coin tosses: $S = \{HH, HT, TH, TT\}, A_1 = \{HH, HT\}, A_2 = \{HH, TH\}, A_3 = \{HT, TH\}$.

A_1 & A_2 are indepedent $P(A_1 \cap A_2) = P(A_1)P(A_2)$

A_2 & A_3 are indepedent $P(A_3 \cap A_2) = P(A_3)P(A_2)$

A_1 & A_3 are indepedent $P(A_1 \cap A_3) = P(A_1)P(A_3)$

A_1, A_2 & A_3 are not still not independent as $P(A_1 \cap A_2 \cap A_3) \neq P(A_1)P(A_2)P(A_3)$.

6 Continuity of Probability

Ω is an equivalent notation for sample space.

Consider the probability space (S, F, P).

$E_i \in F$. Let $E_1 \subseteq E_2 \subseteq \cdots$ be countably infinite sequence of events (increasing sequence of events). Then

$$P(\bigcup_{i=1}^{\infty} E_i) = \lim_{i \to \infty} P(E_i)$$

.

$$\bigcup_{i=1}^{\infty} E_i \in F$$

Pushing the limit from inside the probability expression to outside.

$$P(\bigcup_{i=1}^{\infty} E_i) = \lim_{i \to \infty} P(E_i)$$

$$P(\lim_{n \to \infty} \bigcup_{i=1}^{n} E_i) = \lim_{i \to \infty} P(E_i)$$

Exchanging limits with probability $\rightarrow$ non trivial operation. Exchanging limits with differentiation, integration etc. $\rightarrow$ require a proof.

- $$E_1 \subseteq E_2 \subseteq \cdots$$

  $$P(\bigcup_{i=1}^{\infty} E_i) = \lim_{i \to \infty} P(E_i)$$

  Proof: We shall use the third axiom. $D_1, D_2 \cdots \rightarrow$ mutually exclusive events.

  $$P(\bigcup_{i=1}^{\infty} D_i) = \sum_{i=1}^{\infty} P(D_i)$$

$E_1 \subseteq E_2 \subseteq \rightarrow$ not mutually exclusive.

Hence we construct a sequence which are mututally exclusive and unions of both sequences have to be the same.

$A_1 = E_1,\ A_2 = E_2 \setminus E_1,\ A_3 = E_3 \setminus E_2\ , \cdots.$

$$P(\bigcup_{i=1}^{\infty} E_i) = P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i) = \lim_{n\to\infty} \sum_{i=1}^{\infty} P(A_i)$$

$$\Rightarrow \lim_{n\to\infty} \sum_{i=1}^{n} P(A_i) = \lim_{n\to\infty} P(\bigcup_{i=1}^{n} A_i) = \lim_{n\to\infty} P(E_n)$$

If there exists a subsequence of events, which is increasing.

- Let $E_1 \supseteq E_2 \supseteq \cdots$ be a countably infinte sequence of events(decreasing sequence of events)

$$P(\bigcap_{i=1}^{\infty} E_i) = \lim_{i\to\infty} P(E_i)$$

Proof: We can use the result of the previous proof.

$$E_1^c \subseteq E_2^c \subseteq \cdots$$

which is an increasing set of events

$$P(\bigcup_{i=1}^{\infty} E_i^c) = \lim_{i\to\infty} P(E_i^c)$$

as,

$$\left(\bigcap_{i=1}^{\infty} E_i\right) = \left(\bigcup_{i=1}^{\infty} E_i^c\right)^c$$

$$\Rightarrow 1 - P(\bigcup_{i=1}^{\infty} E_i^c) = \lim_{i\to\infty} 1 - P(E_i^c)$$

$$P\left(\bigcap_{i=1}^{\infty} E_i\right) = \lim_{i\to\infty} P(E_i)$$

7 Random Variables

Consider a probability space (S, F, P) and an experiment is conducted.

Probability spaces vary a lot based on the experiment. We use random variables to be able to develop a theory of probability which is indepedent of the actual experiment which is performed.

$$S \xrightarrow{\text{R.V.}} \mathbb{R}$$
$$F \xrightarrow{\text{R.V.}} B$$

Random Variable is a function which maps the sample space to the real line. It also maps the event space to the Borel σ-algebra, B.

7.1 Borel σ-algebra

It is the smallest σ- algebra which contains sets of the form $(-\infty, x]$ $\forall x \in \mathbb{R}$.

$$B = \{\mathbb{R}, (-\infty, x], (x, \infty), \phi, (-\infty, x), (x, y), [x, y], \{x\}, (x, y], [x, y)\}$$

$$E_i = (-\infty, x_i] \quad x_i = x - \frac{1}{i}$$
$$\bigcup_{i=1}^{\infty}(-\infty, x_i] = \bigcup_{i=1}^{\infty} E_i = (-\infty, x)$$

Random Variable $X : S \to R$

X has to be a measurable function.

A function is said to be measurable if pre-image of $(-\infty, x]$ $\forall x \in \mathbb{R}$ is in the event space.

$$X : S \to R$$
$$\text{inverse image} \leftarrow (-\infty, x]$$
$$X^{-1}((-\infty, x]) \subseteq S \qquad X^{-1}((-\infty, x]) \in F$$

eg: $X : \{1, 2, 3, 4, 5, 6\} \to \{0, 1, 0, 1, 0, 1\}$

$$X^{-}1\left((-\infty, 0.5]\right) = \{1, 3, 5\}$$

In the above example, X is not an invertible mapping. We are using X^{-1} as a notation.

X is a map from S to $\mathbb{R}$ such that

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

$$X^{-1}\left((-\infty, x]\right) \leftarrow (-\infty, x] \qquad X^{-1}\left((-\infty, x]\right) \in F$$

7.2 Examples

- X which is a random variable.

  $S = \{a, b, c\}$, $F = \{\phi, \{a\}, \{b, c\}, S\}$, $X : S \to \mathbb{R}$

  $X(\omega) = 0$, $\omega = a$

  $X(\omega) = 1$, $\omega = b, c$

  1. $(-\infty, x]$, $x < 0$, $X^{-1}\left((-\infty, x]\right) = \phi \in F$
  2. $(-\infty, x]$, $0 \leq x < 1$, $X^{-1}\left((-\infty, x]\right) = \phi \in F$
  3. $(-\infty, x]$, $x \geq 1$, $X^{-1}\left((-\infty, x]\right) = S \in F$

- X which is not a random variable $S = \{a, b, c\}$, $F = \{\phi, \{a\}, \{b, c\}, S\}$, $X : S \to \mathbb{R}$

  $X(\omega) = 0$, $\omega = b$

  $X(\omega) = 1$, $\omega = a, c$

  But $X^{-1}((-\infty, x])$, $0 \leq x < 1 = \{b\}$, $\{b\} \notin F$

  X is not a random variable.

If S is finite & F is a power set of S, then every function is a random variable.

$$X^{-1}((-\infty, x]) \subseteq S$$

F is set of all subsets of S.

7.3 Theorem

Let a probability space be (S, F, P) and X is a random variable $X : S \to \mathbb{R}$. The following conditions hold:

- $X^{-1}((-\infty, x)) \in F$
- $X^{-1}(\{x\}) \in F$
- $X^{-1}((x_1, x_2]) \in F$
- $X^{-1}((x_1, x_2)) \in F$

Proof: Applying the axioms of the event space we want to show the above conditions are true.

$A_i = X^{-1}((-\infty, x])$ $\quad x_i = x - \frac{1}{i}$. Now,

$$\bigcup_{i=1}^{\infty} (-\infty, x_i] = (-\infty, x)$$

$$\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} X^{-1}((-\infty, x_i]) = X^{-1}((-\infty, x)) \in F$$

$$X^{-1}((-\infty, x] \cap (-\infty, x)^c) = X^{-1}(\{x\}) \in F$$

Similarly the other conditions can also be proved.

7.4 Cumulative distribution function

It is denoted by $F_X(x)$.

$$F_X(x) = P(X \leq x) = P((-\infty, x]) = P(X^{-1}((-\infty, x]))$$

The above definitions are different notations used, the last one is well defined from the probability space.

7.5 Properties of CDF

1. $F_X(x)$ is a non-decreasing function of x.

   Proof: If $x_2 \geq x_1$, then $F_X(x_2) \geq F_X(x_1)$.

$$F_X(x_2) = P(X \leq x_2) = P(X \leq x_1) + P(x_1 < X \leq x_2)$$
$$F_X(x_2) = P(X \leq x_2) \geq P(X \leq x_1) - F_X(x_1)$$

2. $\lim_{x\to\infty} F_X(x) = 1$

   Proof: Construct a decreasing set of events. $A_i = (-\infty, i]$, $B_i = X^{-1}(A_i)$.

$$\bigcap_{i=1}^{\infty} A_i = \mathbb{R}$$

$$\bigcap_{i=1}^{\infty} B_i = S$$

   Applying continuity of probability,

$$\lim_{i\to\infty} P(B_i) = P(\bigcup_{i=1}^{\infty} B_i) = P(S) = 0$$
$$\lim_{i\to\infty} P(X \leq i) = \lim_{x\to\infty} P(X \leq x) = \lim_{x\to\infty} F_X(x)$$

   The change from integers to real numbers x in the last step is valid.

3. $\lim_{x\to-\infty} F_X(x) = 0$

   Proof: Construct a decreasing set of events. $A_i = (-\infty, -i]$, $B_i = X^{-1}(A_i)$.

$$B_1 \supseteq B_2 \supseteq \cdots$$

$$\bigcap_{i=1}^{\infty} B_i = \phi$$

$$P(\bigcap_{i=1}^{\infty} B_i) = \lim_{i\to\infty} P(B_i) = 0$$
$$\lim_{i\to\infty} P(B_i) = \lim_{i\to\infty} P((-\infty, i]) = \lim_{x\to\infty} P((-\infty, x]) \to (\text{x = all real nos.})$$
$$\lim_{x\to\infty} P((-\infty, x]) = \lim_{x\to\infty} F_X(x) = 0$$

7.6 Indicator Random Variable

$$I_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases}$$

$$I_A : S \to \mathbb{R}$$

$$F_{I_A}(x) = P(I_A \leq x)$$

If $x < 0$, $P(I_A \leq x) = 0$.
If $0 \leq x < 1$, $P(I_A \leq x) = 1 - P(A)$.
If $x \geq 1$, $P(I_A \leq x) = 1$.