STAT 712 Chapter Read Pre-class

Chapter 1.1

Def 1.1.1 : The set, S, of all possible outcomes of a particular experiment is called the sample space for the experiment

ie: Coin toss $S = \{T, H\}$

- A sample space is countable (discrete) if $\exists$ a 1-1 correspondence between the sample and integers. ie: SAT Score
- A sample space is uncountable (continuous) if $\nexists$ a 1-1 correspondence between the sample and integers ie: time or distance

Def 1.1.2 : An event is any collection of possible outcomes of an experiment, that is, any subset of S

- Let A be an event, a subset of S

if $A \subset B \Leftrightarrow x \in A \Leftrightarrow x \in B$

if $A = B \Leftrightarrow A \subset B \Leftrightarrow B \subset A$

- Union : $A \cup B = \{x \mid x \in A \vee x \in B\}$
- Intersection: $A \cap B = \{x \mid x \in A \wedge x \in B\}$
- Complementation: $A^c = \{x \mid x \in S \wedge x \notin A\}$

example: Choose 1 card from the deck

$S = \{$clubs, hearts, spades, diamonds$\}$

$A = \{$clubs, diamonds$\}$ $B = \{$diamonds, spades, hearts$\}$

$A \cup B = \{c, d, s, h\} = S$ $A \cap B = \{d\}$ $A^c = \{H, S\}$

Theorem 1.1.4

- Commutativity: $A \cup B = B \cup A$ & $A \cap B = B \cap A$
- Associativity: $A \cup (B \cup C) = (A \cup B) \cup C$ & $A \cap (B \cap C) = (A \cap B) \cap C$
- Distributive Laws: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ & $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- DeMorgan's Law: $(A \cup B)^c = A^c \cap B^c$ & $(A \cap B)^c = A^c \cup B^c$

Distributive Proof: $A \cap (B \cup C) = \{x \mid x \in A \wedge x \in (B \cup C)\}$

$(A \cap B) \cup (A \cap C) = \{x \mid x \in (A \cap B) \vee x \in (A \cap C)\}$

$x \in A \wedge x \in (B \cup C) \stackrel{?}{=} x \in (A \cap B) \vee x \in (A \cap C)$

So let $x \in (A \cap (B \cup C))$ $\therefore$ $x \in A \wedge (x \in B \vee x \in C)$ $\therefore$ $[x \in (A \cap B) \vee x \in (A \cap C)]$ $\Rightarrow$

1st part of proof need $\Leftarrow$ for full pf

- Operations on collections: $\bigcup_{i=1}^{\infty} A_i = \{x \mid x \in A_i \text{ for some } i\}$

$$\bigcap_{i=1}^{\infty} A_i = \{x \mid x \in A_i \text{ for all } i\}$$

example: let $S = (0,1]$ and $A_i = [1/i, 1]$ then

$$\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} [(1/i), 1] = \{x \mid x \in [1/i, 1] \text{ for some } i\} = (0,1]$$

$$\bigcap_{i=1}^{\infty} A_i = \bigcap_{i=1}^{\infty} [(1/i), 1] = \{x \mid x \in [1/i, 1] \text{ for all } i\} = \{1\}$$

Def 1.1.5
- A and B are <u>disjoint</u> or <u>mutually exclusive</u> if $A \cap B = \emptyset$
- the events $A_1, A_2, \ldots$ are <u>pairwise disjoint</u> or <u>mutually exclusive</u> if $A_i \cap A_j = \emptyset$ for all $i \neq j$

Def 1.1.6
- If $A_1, A_2, \ldots$ are pairwise disjoint and $\bigcup_i^{\infty} A_i = S$ then the collection forms a <u>partition</u> of S. It divides S into small non-overlapping pieces

<u>Chapter 1.2</u>

Def 1.2.1 <u>Sigma Algebra</u> or <u>Borel field</u> (B) is a collection of subsets of S that:

a) $\emptyset \in B$

b) If $A \in B$ then $A^c \in B$ (closed under complementation)

c) If $A_1, A_2 \ldots \in B$ then $\bigcup_i^{\infty} A_i \in B$ (closed under countable unions)

d) If we use DeMorgan's Law $(\bigcup_i^{\infty} A_i^c)^c = \bigcap_i^{\infty} A_i$ by property b and c

Sigma-Algebra-I $\longrightarrow$ If S is finite or countable then these technicalities do not arise

So $B = \{$ all subsets of S including S itself $\}$

$\therefore$ If S has n elements there are 2^n sets in B

Sigma-Algebra II $\rightarrow$ Let $S = (-\infty, \infty)$, the real line. Then B is chosen to contain all sets of the form $[a,b]$, $[a,b)$, $(a,b]$, (a,b) for all $\mathbb{R}$ a and b

Def 1.2.4 Given a sample space S and an associated sigma algebra $\mathcal{B}$ a probability function is a function P w/ domain $\mathcal{B}$ that satisfies

axioms of probability
(1) $P(A) \geq 0$ for all $A \in \mathcal{B}$
(2) $P(S) = 1$
(3) If $A_1, A_2, \ldots \in \mathcal{B}$ are pairwise disjoint then $P(\bigcup_i^\infty A_i) = \sum_i^\infty P(A_i)$

Sometimes replaces 3 (3a) If $A \in \mathcal{B}$ and $B \in \mathcal{B}$ are disjoint then $P(A \cup B) = P(A) + P(B)$

EXAMPLE $S = \{H, T\}$ "Coin toss"

$P(H) = P(T)$ (1)

Since $S = \{H\} \cup \{T\}$:

$P(\{H\} \cup \{T\}) = 1$ (2)

Also H and T are disjoint, so $P(\{H\} \cup \{T\}) = P(\{H\}) + P(\{T\})$

and $P(H) + P(T) = 1$ (3)

If we solve 1 and 2 together we get $P(H) = P(T) = 2$

Theorem 1.2.6: Let $S = \{s_1, \ldots, s_n\}$ be a finite set. Let $\mathcal{B}$ be any sigma algebra of subsets of S. Let $p_1, \ldots, p_n$ be non negative numbers that sum to 1. For any $A \in \mathcal{B}$ define $P(A)$ by

$$P(A) = \sum_{s_i \in A} p_i$$

P is a probability function on $\mathcal{B}$. This remains true if $S = \{s_1, s_2, \ldots\}$ is a countable set

PROOF: Proof for finite S:

1) for any $A \in \mathcal{B}$, $P(A) = \sum_{\{i : s_i \in A\}} p_i \geq 0$ b/c every $p_i \geq 0$

2) thus $P(S) = \sum_{s_i \in S} p_i = \sum_i^n p_i = 1$

3) Let $A_1, \ldots, A_k$ denote pairwise disjoint events ($\mathcal{B}$ contains only a finite # of sets so we need consider only finite disjoint unions) Then

$$P\left(\bigcup_i^k A_i\right) = \sum_{\{j : s_j \in \bigcup_i^k A_i\}} p_j = \sum_i^k \sum_{\{j : s_j \in A_i\}} p_j = \sum_i^k P(A_i)$$

example: Darts

$P(\text{scoring } i \text{ points}) = \dfrac{\text{Area of region } i}{\text{Area of dartboard}}$

ie $P(\text{scoring } 1 \text{ pt}) = \dfrac{\pi r^2 - (4/5 \cdot r)^2 \pi}{\pi r^2} = 1 - \frac{16}{25} = .36$

ie $P(\text{scoring } i \text{ pts}) = \dfrac{\pi r^2 - ((5-i)/5)^2 \pi}{\pi r^2}$

- The events 1, 2, 3, 4 and 5 are disjoint regions

$\therefore \exists$ 5 outcomes w/ 5 probabilities Σ to 1

$\therefore$ by T 1.26 we have a probability function

Theorem 1.28 If P is a probability function and A is any set in $\mathcal{B}$ then

a) $P(\emptyset) = 0$

b) $P(A) \le 1$

c) $P(A^c) = 1 - P(A)$

Proof: c) $A^c \cup A = S$

$P(S) = 1$

$P(A \cup A^c) = P(A) + P(A^c)$ because A and A^c are disjoint

$\therefore 1 = P(A) + P(A^c)$

$\therefore 1 - P(A) = P(A^c)$ //

b) Since $P(A^c) \ge 0$ b is immediately implied by c

a) $S \cup \emptyset = S$

$P(S) = 1 = P(S \cup \emptyset) = P(S) + P(\emptyset)$

$\therefore P(\emptyset) = 0$

Theorem 1.29 If P is a probability function and A and B are any sets in $\mathcal{B}$

a) $P(B \cap A^c) = P(B) - P(A \cap B)$

b) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

c) If $A \subset B$ then $P(A) \le P(B)$

NOTE: $B = (B \cap A) \cup (B \cap A^c) \ni B \in S$ and $A \in S$

a) $P(B) = P(\{B \cap A\} \cup \{B \cap A^c\}) = P(B \cap A) + P(B \cap A^c)$ by note + 1.24 3a

b) Note: $A \cup B = A \cup \{B \cap A^c\}$

$P(A \cup B) = P(A) + P(B \cap A^c) = P(A) + P(B) - P(A \cap B)$ by note + 1.29a

c) Note: if $A \subset B$ then $A \cap B = A \therefore 0 \le P(B \cap A^c) = P(B) - P(A)$ by 1.29a

• Since $P(A \cup B) \le 1$

$P(A)+P(B)-P(A\cap B) \le 1$ by 1.2.9 b

$P(A)+P(B)-1 \le P(A\cap B)$ by algebra

• This is known as Bonferroni's Inequality - which is useful when it is difficult or even impossible to calculate the intersection probability

• Spse $P(A)=P(B)=.95$

$P(A\cap B) \ge P(A)+P(B)-1 = .9$

WARNING: unless $P(A)\ P(B)$ is large Bonferroni will return a neg., but correct, answer

Theorem 1.2.11 - If P is a probability function

a: $P(A)=\sum_1^\infty P(A\cap C_i)$ for any partition $C_1, C_2, C_3, \dots$

b: $P(\bigcup_1^\infty A_i) \le \sum_1^\infty P(A_i)$ for any sets $A_1, A_2, \dots$ (Boole's Inequality)

Proof: • Since $C_1, C_2, \dots$ form a partition we have that $C_i \cap C_j = \emptyset\ \forall\ i \ne j$ Def 1.1.5 and $S=\bigcup_1^\infty C_i$ by Def 1.1.6 Hence

$$A = A\cap S = A\cap \left(\bigcup_1^\infty C_i\right) = \bigcup_1^\infty (A\cap C_i)$$ by distributive

• B/c $A\cap C_i$ are disjoint and C_i is disjoint we have

(a) $$P\left(\bigcup_1^\infty A\cap C_i\right) = \sum_1^\infty P(A\cap C_i)$$ //

• $A_1^*, A_2^*, \dots$ be a disjoint collection $\ni \bigcup_1^\infty A_i^* = \bigcup_1^\infty A_i$

• We define A_i^* by

$$A_1^* = A_1,\quad A_i^* = \underbrace{A_i \setminus \left(\bigcup_{j}^{i-1} A_j\right)}_{\text{the part of } A_i \text{ not seen in any } A_j \ \ni\ j \le i-1}$$ for $i=2,3,4,\dots$ and $A\setminus B = A\cap B^c$

then $P\left(\bigcup_1^\infty A_i\right) = P\left(\bigcup_1^\infty A_i^*\right) = \sum_{i=1}^\infty P(A_i^*)$ b/c A^* is disjoint

$A_i^* \cap A_k^* = \{A_i \setminus (\bigcup_j^{i-1} A_j)\} \cap \{A_k \setminus (\bigcup_1^{k-1} A_j)\}$ definition of A_i^*

$= \{A_i \cap (\bigcup_j^{i-1} A_j)^c\} \cap \{A_k \cap (\bigcup_1^{k-1} A_j)^c\}$ definition of $\setminus$

$= \{A_i \cap (\bigcap_j^{i-1} A_j^c)\} \cap \{A_k \cap (\bigcap_1^{k-1} A_j^c)\}$ De Morgan's law

So $A_i^* \subseteq A$ so $P(A_i^*) \le P(A_i)$

(b) $$\sum_1^\infty P(A_i^*) \le \sum_1^\infty P(A_i)$$

Boole's Inequality and Bonferroni's Inequality.

Essentially the same:

Bonferroni → $\boxed{P(A \cap B) \geq P(A) + P(B) - 1}$

Apply Boole's to $A^c \Rightarrow P(\bigcup_1^n A_i^c) \leq \sum_1^n P(A_i^c)$

- since $\bigcup A_i^c = (\bigcap A_i)^c$ and $P(A_i^c) = 1 - P(A_i)$ we obtain

$$1 - P(\bigcap_1^n A_i) \leq n - \sum_1^n P(A_i)$$

which becomes: $\boxed{P(\bigcap_{i=1}^n \geq \sum_1^n P(A_i) - (n-1)}$

COUNTING

Lottery - I: for a number of years the NY state lottery was the following
Pick 6 #s 1-44 person w/ matching 6 randomly selected #s wins

$$P(\text{Win}) = \frac{1}{\#\text{ of ways}} = \frac{1}{{}_{44}P_6} \approx 1.9675 \times 10^{-10}$$

$${}_{44}P_6 = \frac{44!}{(44-6)!}$$

Tournament: Single elimination w/ 16 players How many paths to victory for a player

Theorem 1.2.14: If a job consists of k separate tasks, the i^{th} of which can be done in n_i ways, $i = 1, \ldots, k$ then the entire job can be done in $n_1 \times \cdots n_k$ ways

Proof: for $k = 2$: $\underbrace{(1 \times n_2) + (1 \times n_2) + \cdots + (1 \times n_2)}_{n_1 \text{ terms}} = n_2 \cdot n_1$

- the first task can be done n_1 ways for which each has n_2 choices for the second task

example the first 2 balls in the lottery: 44 different balls

first ball: 44 ways
second ball $\times$ 43 ways
1892 ways

Possible methods of counting

	w/o replacement	w/ replacement
Ordered	$_nP_r = \frac{n!}{(n-r)!}$	n^r
Unordered	$_nC_r = \binom{n}{r} = \frac{n!}{r!\,(n-r)!}$	$\binom{n+r-1}{r} = \frac{(n+r-1)!}{r!\,(n-1)!}$

Def 1.2.16: $n! = n \times (n-1) \times \cdots \times 2 \times 1$
$0! = 1$

Def 1.2.17 for nonnegative $n, r \in \mathbb{Z} \ni n \geq r$
$\binom{n}{r} = n$ choose $r = \frac{n!}{r!\,(n-r)!}$

Lottery: ordered w/o replacement

$44 \times 43 \times \cdots \times 39 = \frac{44!}{38!} = 5{,}082{,}517{,}440$

Lottery: Ordered w/ replacement

$44 \times 44 \times \cdots \times 44 = 44^6 = 7{,}256{,}313{,}856$

Lottery: Unordered w/o replacement

$(44!/38!)/6!$ (# of ways to arrange 6 #s) $= \frac{44!}{6!\,38!} = 7{,}059{,}052$

Lottery: Unordered w/ replacement *hardest

1	2	3	4	...	41	42	43	44
M		MM		...	M	M		M

- Think of it a 49 things: 43 | and 6M
- We can order these 49! ways
- $49!/\underbrace{6!\,43!}_{\text{We must divide by } 6!\,43! \text{ to remove redundents}} = 13{,}983{,}816$ ways

- When S is a finite set $\ni$ all outcomes are equally likely

$$P(A) = \sum_{s_i \in A} P(\{s_i\}) = \sum_{s_i \in A} \frac{1}{N} = \frac{\#\text{ of elements in } A}{\#\text{ of elements in } S}$$

Example: Consider choosing a 5 card poker hand from a 52 card deck

$\exists \binom{52}{5} = \frac{52!}{5!\,47!} = 2,598,960$

So $P(\text{4 Aces}) = \frac{1 \cdot (52-4)}{2,598,960}$

So $P(\text{4 of a kind}) = \frac{13\,(52-4)}{2,598,900}$

So $P(\text{Exactly 1 pair}) = \underbrace{13}_{\text{Denoms to chose from}} \cdot \underbrace{\binom{4}{2}}_{\text{ways to pick 2 cards w/ same denom}} \cdot \underbrace{\binom{12}{3}}_{\text{ways to pick 3 other cards}} \cdot \underbrace{4^3}_{\text{ways to pick the 3 other cards}} = 1,098,240$

$\doteq \frac{1098240}{2,598,960}$

- Sample 2 from 3 w/ replacemet

unordered	$\{1,1\}$	$\{2,2\}$	$\{3,3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$
ordered	$(1,1)$	$(2,2)$	$(2,2)$	$(1,2)(2,1)$	$(1,3)(3,1)$	$(2,3)(3,2)$
	$1/9$	$1/9$	$1/9$	$2/9$	$2/9$	$2/9$

common misconception: $P = 1/6$ based on unordered pairs

- Calculate an average of 4 #'s chosen w/ replacement: 2, 4, 9, 12
  - because we are looking at average order doesn't matter
  - unordered w/ replacement $\binom{n+n-1}{n}$ b/c $n=r$ — ordered w replacement $(4)^4 = 256$
  - avg = 4.75 can occur only if our set contains one 2, two 4s, and one 9

unordered: $\{2,4,4,9\}$

ordered: $4!/2$ (ways to order $\{2,4,4,9\} = 24/2 = 12$ b/c of the 4s) 2494_2 vs 24_294

$(2,4,4,9)\ (2,4,9,4)$

$(2,9,4,4)\ (4,2,4,9) \ldots$

- $P(\{2,4,4,9\}) \overset{\text{CORRECT}}{=} 12/256$ vs $1/\overset{\text{INCORRECT}}{\binom{4+4-1}{4}} = 1/35$

1.3

We know → • Chances 4 cards selected from a deck of cards are all aces $1/\binom{52}{4}$

What about → • Chances 4 cards selected are aces, given i aces in i cards occurred $= P(4|i)$

$P(4|i) = P(4 \cap i)/P(i) = P(4 \cap i)/\left[\binom{4}{i}/\binom{52}{i}\right] = \binom{52}{i}/\binom{4}{i}\binom{52}{?}$

• Conditional Probability is tricky so it must be thought of carefully

• Three Prisoner's: • Warden says he will pardon one at random ∴ each prisoner has a 1/3 chance of receiving a pardon

• The warden tells A that B will be executed, A thinks that his chances have gone up to 1/2

Let W denote that the warden says B will die → $P(A|W) = \frac{P(A \cap W)}{P(W)}$

(22) $P(W) = P(B \text{ dies} + A \text{ pardon}) + P(B \text{ dies} + C \text{ pardon}) + P(B \text{ dies} + B \text{ pardon})$

$\underbrace{\frac{1}{6} + \frac{1}{3}}_{\text{why are these different?}} + 0$

$P(A|W) = P(A \cap W)/P(W)$

Theorem 1.3.5 Bayes' Rule: Let $A_1, A_2, \ldots$ be a partition of the sample space and let B be any set. Then $\forall\ i = 1, 2, 3, \ldots$

$$P(A_i|B) = \frac{P(B|A_i)\,P(A_i)}{\sum_{j}^{\infty} P(B|A_j)\,P(A_j)}$$

Example: • $P(\text{dot sent}) = 3/7$ $\quad P(\text{dash sent}) = 4/7$ $\quad$ Morse code $\overset{\text{dot}}{3} : \overset{\text{dash}}{4}$

• $P(\text{dot sent}|\text{dot received}) = P(\text{dot received}|\text{dot sent}) \frac{P(\text{dot sent})}{P(\text{dot received})}$

• $P(\text{dot received}) = P(\text{dot received} \cap \text{dot sent}) + P(\text{dot received} \cap \text{dash sent}) = [P(\text{dot received}|\text{dot sent}) \times P(\text{dot sent})] + [P(\text{dot received}|\text{dash sent})\,P(\text{dash sent})]$

$= 7/8 \times 3/7 + 1/8 \times 4/7 = 25/56$

∴ $P(\text{dot sent}|\text{dot received}) = ((7/8)(3/7))/(25/56) = \frac{21}{25}$

definition $P(A|B) = P(A)$ if A & B are independent

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

Def 1.3.7 A and B are statistically independent: $P(A \cap B) = P(A)P(B)$

EXAMPLE 1.3.8 $P(1 \text{ six in 4 dice rolls}) = 1 - P(\text{no 6 in 4 rolls}) = 1 - (5/6)^4 = .514$

Theorem 1.3.9 If A and B are independent then the following are independent

a. A and B^c

b. A^c and B

c. A^c and B^c

Proof of a: WTS: $P(A \cap B^c) = P(A)P(B^c)$

$$
\begin{aligned}
P(A \cap B^c) &= P(A) - P(A \cap B) && \text{by 1.2.9a} \\
&= P(A) - P(A)P(B) && A \text{ \& } B \text{ are independent} \\
&= P(A)(1 - P(A)) && \text{Dist} \\
&= P(A)P(B^c) \quad /\!/ \text{EMS}
\end{aligned}
$$

Ex 1.3.10 S = All possible outcomes of 2 dice being rolled $n = 6^2 = 36$

A = Rolling Doubles

B = Sum of dice between 7 and 10

C = sum is 2, 7, or 8

(1,1) (2,2) (3,3) (4,4) (5,5) (6,6)

$P(A) = \frac{6}{36} = \frac{1}{6}$

(3,4) (3,5) (3,6) (4,4) (5,4) (6,4) (5,2) (6,1) (6,2)
(4,3) (5,3) (6,3) (4,5) (4,6) (5,5) (2,5) (1,6) (2,6)

$P(B) = 18/36 = 1/2$

(1,1) | (1,6) (2,5) (3,4) | (2,6) (3,5) (4,4)
| (6,1) (5,2) (4,3) | (6,2) (5,3)

$P(C) = 1/36 + 6/36 + 5/36 = 1/3$

$P(A \cap B \cap C) = P(A)P(B)P(C) = 1/36$ $\Rightarrow A \cap B \cap C$ = sum is 8 w/ double 4s

But $P(B \cap C) = 11/36 \neq P(B)P(C)$

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

Def 1.3.12 $A_1, \dots, A_n$ are mutually independant for any sub collection $A_{i_1}, \dots A_{i_k}$ we have:

$$P\left(\bigcap_{1}^{k} A_{i_j}\right) = \prod_{1}^{k} P(A_{i_j})$$

Ex 1.3.13 Tossing a coin in 3 tosses $\quad 2^3 = 8$ outcomes

Let H_i for $i = 1,2,3$ denote the event that the i^{th} toss is a head

ie $H_1 = \{HHH, HHT, HTH, HTT\}$ if it's a fair coin $P(H_1) = 1/2 = P(H_2) = P(H_3)$

since H_1, H_2, H_3 are mutually independent

$$P(H_1 \cap H_2 \cap H_3) = (1/2)^3 = 1/8 = P(H_1)\cdot P(H_2)\cdot P(H_3)$$

$$P(H_1 \cap H_2) = (1/2)^2 = 1/4 = P(H_1)\cdot P(H_2)$$

$\boxed{1.4}$

<u>Random Variables</u> a function from a sample space S into the real #'s

EXPERIMENT	Random Variable
Toss 2 dice	X = sum of 2 dice
Toss a coin 25 times	X = # of heads in 25 tosses
Apply different amt of fertilizer	X = yield/yard

So we go from $S = \{s_1, s_2, s_3 \dots s_n\}$ w/ Prob. function P to

$X = \{x_1, \dots, x_m\}$ w/ $P_X(X = x_i) = P\{s_j \mid X(s_j) = x_i\}$

We call P_X = an induced probability function on X

IDEA: RV abstracts observation to one number $\in \mathbb{R}$

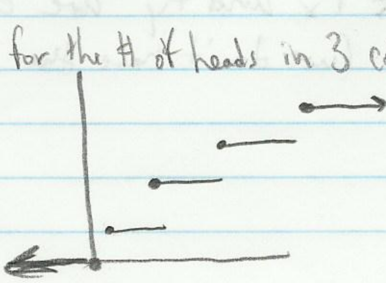
P_X for uncountable $x \Rightarrow P_X(X \in A) = P(\{s \in S \mid X(s) \in A\})$

1.5

<u>Cumulative distribution function</u> - $F_X(x) = P_X(X \le x)\ \forall x$ Return the probability that $X \le x$

example $F_X(x) = \begin{cases} 0 & \text{if } -\infty < x < 0 \\ 1/8 & 0 \le x < 1 \\ 1/2 & 1 \le x < 2 \\ 7/8 & 2 \le x < 3 \\ 1 & 3 \le x < 4 \end{cases}$ for the # of heads in 3 coin tosses

Note $F_X(x)$ defined on all x



Theorem 1.5.3 $F(x)$ is a cdf iff

a) $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$

b) $F(x)$ is a non-decreasing function of x

c) $F(x)$ is right-continuous; that is $\forall x_0$ $\lim_{x\to x_0^+} F(x) = F(x_0)$

Proof a) $\lim_{x\to-\infty} F(x) = \lim_{x\to-\infty} P(X \le x) = 0$

$\lim_{x\to\infty} F(x) = \lim_{x\to\infty} P(X \le x) = 1$

b) If $x_0 \le x_1$ then $F(x_0) = P(X \le x_0) \le P(X \le x_1) = F(x_1)$

$\therefore$ $F(x)$ is non-decreasing

c) $F(x) = P(X \le x)$ is right-continuous because even if the graph is discontinuous because we define $F(x)$ using '$\le$' $\lim_{x\to x_0^+} F(x) = F(x_0)$ //

Example: $P_x(X = x) = \underbrace{(1-p)^{x-1}}_{\text{Prob of tails all other trials}} \underbrace{p}_{\text{Prob of heads}}$ = Prob of a head on any given toss

So $P(X \le x) = \sum_1^x P(X = i) = \sum_1^x (1-p)^{i-1} p$

Def 1.5.7 A random variable is cont if $F_X(x)$ is continuous and discrete if it is a step function

ie $F_X(x) = \frac{1}{1+e^{-x}}$ is continuous $\qquad f_X(x) = \begin{cases} \frac{1-\epsilon}{1+e^{-x}} & x < 0 \\ \epsilon + \frac{1-\epsilon}{1+e^{-x}} & x \ge 0 \end{cases}$ is discrete

$\lim_{x\to\infty} = 0$

$\lim_{x\to-\infty} = 1$

Def 1.5.8 The random variables X and Y are identically distributed if for every set $A \in \mathcal{B}^1$, $P(X \in A) = P(Y \in A)$ where $X = Y$ OR $X \ne Y$ $\quad (X \stackrel{d}{=} Y)$ (so $X \stackrel{d}{=} Y \not\Rightarrow X = Y$)

ie X = # of heads observed and Y = number of tails observed : Toss 3 times

Easily shown $P(Y \le y) = P(X \le x)$

Theorem 1.5.10 $F_X(x) = F_Y(y) \rightarrow X$ and Y are identically distributed $(X \stackrel{d}{=} Y)$

Proof $\Rightarrow$ $F_X(x) = P(X \le x) = P(Y \le y) = F_Y(x)$ $\forall x$ $P(X \in A), P(Y \in A)$

$\Leftarrow$ F_X and F_Y are equal on all intervals WTS on all sets

(we don't have the tools yet

$\boxed{1.6}$

Def 1.6.1 The Probability mass function (pmf) of a discrete random variable X is given by $f_X(x) = P(X=x) \ \forall x$

ie $f_X(x) = P(X=x) = \begin{cases} (1-p)^{x-1}p & \text{for } x = 1, 2, \ldots n \\ 0 & \text{otherwise} \end{cases}$

$\vdots$

$P(a \le X \le b) = \sum_a^b f_X(k)$

$P(X \le b = \sum_1^b f_X(k)$

Def 1.6.3 The probability density function or pdf, $f_X(x)$, of a continuous random variable X is given by $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt \ \forall x = P(X \le x)$

$\vdots$

$P(a \le X \le b) = \int_a^b f_X(t)\,dt$

$P(X = x_0) = 0 \quad \forall x_0$

Theorem 1.6.5: A function $f_X(x)$ is a pdf or pmf for x iff

a: $f_X(x) \ge 0 \ \forall x$

b: $\sum_x f_x = 1$ OR $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$

(PMF) (PDF)

$\boxed{\text{Proof}}$ $\Rightarrow$) pdf $1 = \lim_{x\to\infty} F_X(x) = \int_{-\infty}^{\infty} f_X(t)\,dt$ (by def)

pmf $1 = \lim_{x\to\infty} F_X(x) = \sum_{-\infty}^{\infty} f_X(k)$ "

pdf $\int_a^b f_X(t)\,dt \ge 0 \ \forall a,b \in \mathbb{R}$ "

pmf $\sum_a^b f_X(k) \ge 0 \ \forall a,b \in \mathbb{Z}$ "

$\Leftarrow$ 1) Let $F_X(x) = \sum_a^x f_X$

2) Let $F_X(x) = \int_a^b f_X(x)\,dx$

$\therefore$ by 1.5.3 $F_X(x)$ is a pdf or pmf