Lecture 5,

September 3

HW 1: this Friday on Canvas (09/06/24)

Due date: (09/13/24)

TA Office hours: TR 10:50 -- 11:50

Random variables

Probability functions for random variables

- We have $P(A)$ defined on events $A \subset S$, which can be used to assign probabilities for events concerning a r.v. X on $\mathbb{R}$ $(X : S \to \mathbb{R})$

$P: S^* \longrightarrow [0,1]$, $\forall A \subset S$, $P(A) \in [0,1]$

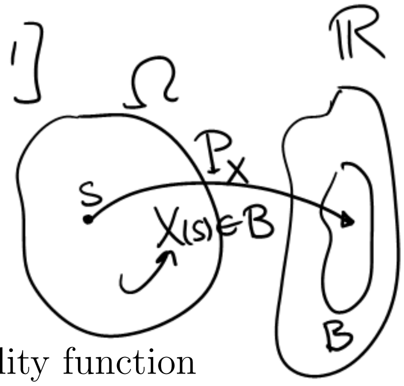
S^* is a set including all possible events in S

- Define $P_X(\cdot)$ for events $B \subset \mathbb{R}$ as follows:

define $B \subseteq \mathbb{R}$

$$P_X(B) = P_X(X \in B) = P(\{s \in S : X(s) \in B\})$$

$\longrightarrow P_X(B) \in [0,1]$



- $P_X(\cdot)$ satisfies the axioms and is therefore a legitimate probability function

$B \subset \mathbb{R}$, e.g. $B = [a,b]$ $P_X(B) = P_X([a,b]) = P(s : a \le X(s) \le b)$

P_X is defined on subsets of $\mathbb{R}$

P is defined on subsets of S

Example 1: A sample space S is generated by flipping a fair coin 4 times. Let $X = \#$ of tails in four flips.

16 outcomes in S

$S = \{TTTT, HHHH, \ldots \text{etc}\}$

Range of $X(s)$, $s \in S$ is $\{0,1,2,3,4\}$

$P_X(X=1) = P(s \in S, X(s) = 1\} = P(\{THHH, HTHH, HHTH, HHHT\})$

$B = \{1\}$, $P_X(X=1) = P_X(B) = P_X(X \in B) = 4/16$

$P_X(X \le 1.5) = P_X(X=0 \text{ or } X=1)$

$= P_X(X=0) + P_X(X=1) = \boxed{5/16}$

$P(\{HHHH\})$ 1/16 4/16

Random variables

Probability functions for random variables (cont'd)

Example 2: Consider an experiment where 10 people $(A, B, C, D, E, F, G, H, I, J)$ are randomly seated (arranged) around a circular table with 10 seats.

(Here the natural space S' is the set of $10!$ arrangements of people at table)

$X(\text{arrangement}) = \#$ of seats between A & B

range of $X \in \{0, 1, 2, 3, 4\}$

$$P_X(X=0) = \frac{\binom{10}{1}(2)\,8!}{10!} = 2/9$$

Pick 1 seat for A

Choices for B

arrangements for others

$$P_X(X=1) = P_X(X=2) = P_X(X=3) = 2/9$$

$$P_X(X=4) = 1/9 = \frac{\binom{10}{1} \times 1! \times 8!}{10!}$$

- Note P and P_X are defined on different probability spaces
  - $P(\cdot)$ is again defined on subsets of S
  - $P_X(\cdot)$ is defined on subsets of $\mathbb{R}$, i.e., $P_X(B) = P_X(X \in B)$ for $B \subset \mathbb{R}$

- It is also convenient (even if slightly sloppy) to write $P(X \in A)$ for $A \subset \mathbb{R}$ (without the subscript X)

Random variables

Cumulative distribution function (cdf)

Definition: The **cumulative distribution function (cdf)** of a random variable X, denoted by $F(\cdot)$, is defined by

$$F(x) = P(X \leq x), \quad \text{any } x \in \mathbb{R}$$

$$= P_X(X \leq x) = P_X(X \in B) = P_X(X \in (-\infty, x])$$

Sometimes written with subscript $F_X(x)$

A function $F(x)$, $x \in \mathbb{R}$, is a cdf for some random variable **if and only if** the following hold:

$F : \mathbb{R} \longrightarrow [0,1]$

1. $F(x)$ is a nondecreasing function of x
2. $\lim_{x \to -\infty} F(x) = 0 \qquad \lim_{x \to \infty} F(x) = 1$

   $\lim_{x \to +\infty} F(x) = \lim_{x \to +\infty} P(X \leq x) = 1$
3. $F(x)$ is right continuous, i.e., $\lim_{x \downarrow x_0} F(x) = F(x_0)$ for any $x_0 \in \mathbb{R}$

e.g., to show part 1., take $y \geq x$ and $F(y) - F(x) = P(X \leq y) - P(X \leq x)$.

$\{X \leq x\} \subset \{X \leq y\} \Rightarrow F_X(x) = P(\{X \leq x\}) \leq P(\{X \leq y\}) = F_X(y)$

$A \subset B \Rightarrow P(A) \leq P(B)$

Example of a cdf: 10 people seated randomly around a table; $X = \#$ of seats between $A \& B$

$X \in \{0, 1, 2, 3, 4\}$

$$\begin{cases} P(X=0) = P(X=1) = P(X=2) = P(X=3) = 2/9 \\ P(X=4) = 1/9 \end{cases}$$

$F(x) = P(X \leq x), \ \forall x \in \mathbb{R}$, $F(-10) = 0$, $F(5.1) = 1$

$F(1.5) = P(X \leq 1.5) = 4/9$