1

08/26/2020

Lecture 1

- A discrete random variable (r.v) has probability mass function (PMF)

$$P(x) := P(X = x) \quad \text{and the r.v} \quad X \sim P(x) \quad \text{where } x \text{ is the}$$

"realized Value"

- The Cumulative distribution function (CDF) is $F(X) := P(X \leq x)$ and the complementary C.D.F or "survival function is $S(X) = 1 - F(X)$

This r.v has "Support" given by $\text{Support}[X] := \{x : p(x) > 0, x \in \mathbb{R}\}$

$|\text{supp}[X]| \leq |\mathbb{N}|$ Countably infinite at most. Sets of this size are called "discrete" sets.

$\uparrow$ # element in a set

- The support and the PMF are related by the following identity:

$$\sum_{x \in \text{Supp}[X]} P(x) = 1$$

- The most "foundamental" r.v is the Bernoulli:

$$X \sim \text{Bern}(P) := \underbrace{P^x (1-P)^{1-x}}_{P(x)} \quad \text{with} \quad \text{supp}[X] = \{0, 1\}$$

- Let defined the "indicator function"

$$\mathbb{1}_A := \begin{cases} 1 & \text{if } A \\ 0 & \text{if } A^c \end{cases} \Rightarrow X \sim \text{Bern}(P) := P^x (1-P)^{1-x} \, \mathbb{1}_{x \in \{0,1\}}$$

$$\Rightarrow \sum_{x \in \mathbb{R}} P(x) = 1$$

What if $P = 1$?

$$X \sim \text{Bern}(1) = 1^x 0^{1-x} \, \mathbb{1}_{x \in \{0,1\}} = \{1 \text{ w.p } 1 = \mathbb{1}_{x=1}\}$$

$X \sim Deg(1) = \{ 1 \text{ wp } 1 \quad X \sim Deg(c) := \mathbb{1}_{x=c} \}$

↑ degenerate

$X \sim Bern(0) = Deg(0)$

- The Convention in this class is that parameter value (P is the parameter of the Bernoulli) that yield degenerate rv's are not part of the legal "parameter space".

$-p \in (0,1)$

- If we have more than one r.v $x_1, x_2, \dots, x_n$ we can gruope them together in column vector:

$\vec{X} = [x_1, x_2, \dots, x_n]^T$ and ~~we~~ define the "joint mass function" (JMF) as:

$$P_{\vec{X}}(\vec{x}) = P_{x_1, \dots, x_n}(x_1, \dots, x_n) \quad \text{Valid for } \vec{x} \in \mathbb{R}^n \text{ and } \sum_{\vec{x} \in \mathbb{R}^n} P(\vec{x}) = 1$$

$$\forall x \cdot P_{X_1}^{(x)} = P_{X_2}^{(x)} = \dots = P_{X_n}^{(x)}$$

- If $X_1 \overset{d}{=} X_2 \overset{d}{=} \dots \overset{d}{=} X_n$ this denotes "equal in distribution" meaning their PMF's are the same. However, this offers no ~~such~~ simplification of the JMF unless...

$X_1, X_2, \dots, X_n \overset{iid}{\sim}$ that means independent and identically distributed.

$$\Rightarrow \quad P_{\vec{X}}(\vec{x}) = \prod_{i=1}^{n} P(x_i)$$

Let $x_1, x_2 \overset{iid}{\sim} Bern(P)$, let $T_2 = f(x_1, x_2) = x_1 + x_2 \sim P_T(t)$ denote

$$P_T(t) = P_{X_1}^{(x)} * P_{X_2}^{(x)}$$

↑ Convolution operator

$Supp[T_2] = \{0, 1, 2\}$